

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
3 July 2003 (03.07.2003)

PCT

(10) International Publication Number  
**WO 03/054152 A2**(51) International Patent Classification<sup>7</sup>: **C12N**

(21) International Application Number: PCT/US02/39555

(22) International Filing Date:  
10 December 2002 (10.12.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/339,739	10 December 2001 (10.12.2001)	US
60/339,453	11 December 2001 (11.12.2001)	US
60/365,091	14 March 2002 (14.03.2002)	US
60/365,384	14 March 2002 (14.03.2002)	US
60/372,381	12 April 2002 (12.04.2002)	US
60/372,615	12 April 2002 (12.04.2002)	US
10/128,558	22 April 2002 (22.04.2002)	US
60/376,045	24 April 2002 (24.04.2002)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/339,739 (CIP)
Filed on	10 December 2001 (10.12.2001)
US	10/128,558 (CIP)
Filed on	22 April 2002 (22.04.2002)
US	60/339,453 (CIP)
Filed on	11 December 2001 (11.12.2001)
US	60/365,384 (CIP)
Filed on	14 March 2002 (14.03.2002)
US	60/365,091 (CIP)
Filed on	14 March 2002 (14.03.2002)
US	60/372,615 (CIP)
Filed on	12 April 2002 (12.04.2002)
US	60/376,045 (CIP)
Filed on	24 April 2002 (24.04.2002)
US	60/372,381 (CIP)
Filed on	12 April 2002 (12.04.2002)

(71) Applicant (for all designated States except US): **HYSEQ, INC.** [US/US]; 675 Almanor Avenue, Sunnyvale, CA 94085 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **TANG, Y., Tom** [—/US]; 4230 Ranwic Court, San Jose, CA 95118 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **GOODRICH, Ryle, W.** [US/US]; 3461 Saint Susan Place, Los Angeles, CA 90066 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court,

Cupertino, CA 95014 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **WANG, Jian-Rui** [CN/US]; 744 Stendal Lane, Cupertino, CA 95014 (US). **GHOSH, Malabika** [IN/US]; 620 Iris Avenue, Apt. 133, Sunnyvale, CA 94086 (US). **XUE, Aidong, J.** [US/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). **WEHRMAN, Tom** [US/US]; CCSR Mol Pharm 3210, 269 W. Campus Dr, Stanford, CA 94305 (US). **WENG, Gezhi** [US/US]; 415 Moraga Avenue, Piedmont, CA 94611 (US). **ZHOU, Ping** [US/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). **DRMANAC, Radoje, T.** [US/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US). **WANG, Zhiwei** [CN/US]; 836 Alturas Ave #B36, Sunnyvale, CA 94085 (US). **MA, Yunqing** [CN/US]; 280 W. California Ave #206, Sunnyvale, CA 94086 (US). **WANG, Dunrui** [CN/US]; 12252 Pepper Tree Lane, Poway, CA 92064 (US). **CHEN, Rui-hong** [US/US]; 1031 Flying Fish Street, Foster City, CA 94404 (US). **XU, Chongjun** [US/US]; 4918 Manitoba Drive, San Jose, CA 95130 (US). **BOYLE, Bryan, J.** [US/US]; 1947 10th Avenue, San Francisco, CA 94116 (US).

(74) Agents: **RIN-LAURES, Li-Hsien** et al.; Hyseq, Inc., 670 Almanor Avenue, Sunnyvale, CA 94085 (US).(81) Designated States (*national*): AF, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

[Continued on next page]

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.



WO 03/054152 A2



---

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



## NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

### 1. CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part application of U.S. Provisional Application  
5 Serial No. 60/339,739 filed December 10, 2001 entitled "Novel Nucleic Acids and Secreted  
Polypeptides", Attorney Docket No. 811; U.S. Application Serial No. 10/128,558 filed April  
22, 2002 entitled "Novel Nucleic Acids and Polypeptides", Attorney Docket No. 812A,  
which in turn claims the benefit of U.S. Provisional Application Serial No. 60/339,453 filed  
December 11, 2001 entitled "Novel Nucleic Acids and Polypeptides", Attorney Docket No.  
10 812; U.S. Provisional Application Serial No. 60/365,384 filed March 14, 2002 entitled  
"Novel Nucleic Acids and Secreted Polypeptides", Attorney Docket No. 814; U.S.  
Provisional Application Serial No. 60/365,091 filed March 14, 2002 entitled "Novel Nucleic  
Acids and Polypeptides", Attorney Docket No. 815; U.S. Provisional Application Serial No.  
60/372,615 filed April 12, 2002 entitled "Novel Nucleic Acids and Secreted Polypeptides",  
15 Attorney Docket No. 817; U.S. Provisional Application Serial No. 60/376,045 filed April 24,  
2002 entitled "Novel Nucleic Acids and Secreted Polypeptides", Attorney Docket No.  
817CIP; U.S. Provisional Application Serial No. 60/372,381 filed April 12, 2002 entitled  
"Novel Nucleic Acids and Polypeptides", Attorney Docket No. 818; PCT Application Serial  
No. PCT/US00/35017 filed December 22, 2000 entitled "Novel Contigs Obtained from  
20 Various Libraries", Attorney Docket No. 784CIP3A/PCT, which in turn is a continuation-in-  
part application of U.S. Application Serial No. 09/552,317 filed April 25, 2000 entitled  
"Novel Contigs Obtained from Various Libraries", Attorney Docket No. 784CIP, which in  
turn is a continuation-in-part application of U.S. Application Serial No. 09/488,725 filed  
January 21, 2000 entitled "Novel Contigs Obtained from Various Libraries", Attorney  
25 Docket No. 784; PCT Application Serial No. PCT/US01/02623 filed January 25, 2001  
entitled "Novel Contigs Obtained from Various Libraries", Attorney Docket No.  
785CIP3/PCT, which in turn is a continuation-in-part application of U.S. Application Serial  
No. 09/491,404 filed January 25, 2000 entitled "Novel Contigs Obtained from Various  
Libraries", Attorney Docket No. 785; PCT Application Serial No. PCT/US01/03800 filed  
30 February 5, 2001 entitled "Novel Contigs Obtained from Various Libraries", Attorney  
Docket No. 787CIP3/PCT, which in turn is a continuation-in-part application of U.S.  
Application Serial No. 09/560,875 filed April 27, 2000 entitled "Novel Contigs Obtained  
from Various Libraries", Attorney Docket No. 787CIP, which in turn is a continuation-in-

part application of U.S. Application Serial No. 09/496,914 filed February 03, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 787; PCT Application Serial No. PCT/US01/04927 filed February 26, 2001 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 788CIP3/PCT, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/577,409 filed May 18, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 788CIP, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/515,126 filed February 28, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 788; PCT Application Serial No. PCT/US01/04941 filed March 5, 2001 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 789CIP3/PCT, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/574,454 filed May 19, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 789CIP, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/519,705 filed March 07, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 789; PCT Application Serial No. PCT/US01/08631 filed March 30, 2001 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 790CIP3/PCT, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/649,167 filed August 23, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 790CIP, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/540,217 filed March 31, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 790; PCT Application Serial No. PCT/US01/08656 filed April 18, 2001 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 791CIP3/PCT, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/770,160 filed January 26, 2001 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 791CIP, which is in turn a continuation-in-part application of U.S. Application Serial No. 09/552,929 filed April 18, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 791; and PCT Application Serial No. PCT/US01/14827 filed May 16, 2001 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 792CIP3/PCT, which in turn is a continuation-in-part application of U.S. Application Serial No. 09/577,408 filed May 18, 2000 entitled “Novel Contigs Obtained from Various Libraries”, Attorney Docket No. 792; all of which are incorporated herein by reference in their entirety, specifically including, but not limited to, incorporation by reference of the

tables in each application displaying sequence information, ematrix signatures, pfam signatures, signal peptide information, transmembrane domain information, chromosomal localization and tissue distribution information, and/or 3-dimensional structural information.

## 5 2. BACKGROUND OF THE INVENTION

### 2.1 TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

10

### 2.2 BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, circulating soluble factors, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

25

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

30

## 3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules,

cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

5           The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

10           The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-911, or 1823-2478 and are provided in  
15           the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases or unknown. In the amino acids provided in the Sequence Listing, \* corresponds to the stop codon.

20           The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1-911, or 1823-2478 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO: 1-911, or 1823-2478. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO: 1-911, or 1823-2478 or a  
25           degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

30           The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-911, or 1823-2478. The sequence information can be a segment of any one of SEQ ID NO: 1-911, or 1823-2478 that uniquely identifies or represents the sequence information of SEQ ID NO: 1-911, or 1823-2478.

          A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information are

provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-911, or 1823-2478 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-911, or 1823-2478 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., *Science* 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO: 1-911, or 1823-2478; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO: 1-911, or 1823-2478; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-911, or 1823-2478. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO: 1-911, or 1823-2478; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in SEQ ID NO: 1-911, or 1823-2478; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homologue (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in SEQ ID NO: 912-1822, or 2479-3134, or Tables 3A, 3B, 5, or 6.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention (SEQ ID NO: 912-1822, or 2479-3134) also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO: 1-911, or 1823-2478; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such processes is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., *in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and

exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a  
5 polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical  
10 condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant  
15 protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions.  
20 The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a  
25 method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or  
30 monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein.

5 Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives  
10 expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound that binds to a polypeptide of the invention is identified.

The methods of the invention also provide methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals  
15 exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can affect such modulation either on the level of target gene/protein expression or target protein activity.

20 The polypeptides of the present invention (*e.g.* SEQ ID NO: 912-1822, or 2479-3134) and the polynucleotides encoding them (*e.g.* SEQ ID NO: 1-911, or 1823-2478) are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Tables 2A and 2B); for which they have a signature region (as set forth in Tables 3A and 3B); or for which they have  
25 homology to a gene family (as set forth in Tables 4A and 4B). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

## 4. DETAILED DESCRIPTION OF THE INVENTION

### 4.1 DEFINITIONS

30 It must be noted that as used herein and in the appended claims, the singular forms “a”, “an” and “the” include plural references unless the context clearly dictates otherwise.



The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule.

5 Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of  
10 secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only certain portion(s) of the nucleic acids bind or it  
15 may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ  
20 line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source  
25 from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides  
30 which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences

(inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or  
5 "oligonucleotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and  
10 N is A, C, G, or T (U) or unknown. It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is  
15 capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7  
20 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to  
25 about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each  
30 polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NO: 1-911, or 1823-2478.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal

DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-911, or 1823-2478. The sequence information can be a segment of any one of SEQ ID NO: 1-911, or 1823-2478 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO: 1-911, or 1823-2478, or those segments identified in Tables 3A, 3B, 5, or 6. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because  $4^{20}$  possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ( $1 \div 4^{25}$ ) times the increased probability for mismatch at each nucleotide position ( $3 \times 25$ ). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence.

While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5 The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

20 The term "translated protein coding portion" means a sequence which encodes for the full-length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

30 The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or

substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e.g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such

alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (*e.g.*, nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (*e.g.*, yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g.*, *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or

enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell.

- 5 Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

10 The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or  
15 elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

20 The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are  
25 also intended to include proteins containing non-typical signal sequences (*e.g.* Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2): 134 -143) and factors released from damaged cells (*e.g.* Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

30 Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent  
5 conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligonucleotides), 55°C (for 20-  
10 base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" or "substantially similar" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and  
15 subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have  
20 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than  
25 10% (90% sequence identity) and in a further variation of this embodiment, by no more than 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% sequence identity, more preferably at least  
30 98% sequence identity, and most preferably at least 99% sequence identity. Substantially equivalent nucleotide sequence of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, the nucleotide sequence has at least about 65% identity, more preferably at least



about 75% identity, more preferably at least about 80% sequence identity, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least about 95% sequence identity, more preferably at least 98% sequence identity, and most preferably at least 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a new stop codon) should be disregarded. Sequence identity may be determined, *e.g.*, using the Jotun Hein method (Hein, J. (1990) *Methods Enzymol.* 183:626-645). Identity between sequences can also be determined by other methods known in the art, *e.g.* by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

#### 4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1-911, or 1823-2478; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO: 1-911, or 1823-2478; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polynucleotides of any one of SEQ ID NO: 1-911, or 1823-2478. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1-911, or 1823-2478; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing, ; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homologue of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 912-1822, or 2479-3134 (for example, as set forth in Tables 3A, 3B, 5, or 6). Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include entire coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1-911, or 1823-2478 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1-911, or 1823-2478 or a portion thereof as a probe. Alternatively, the polynucleotides of

SEQ ID NO: 1-911, or 1823-2478 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99% sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1-911, or 1823-2478, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, *e.g.* 15, 17, or 20 nucleotides or more that are selective for (*i.e.* specifically hybridize to) any one of the polynucleotides of the invention are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1-911, or 1823-2478, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO: 1-911, or 1823-2478 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology results for the nucleic acids of the present invention, including SEQ ID NO: 1-911, or 1823-2478 can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST (Basic Local Alignment Search Tool) program is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and  
5 Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using FASTXY algorithm may be performed.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a  
10 suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

15 The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic  
20 acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative  
25 choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal  
30 fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for

intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention could be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such

polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature  
5 protein coding sequences corresponding to any one of SEQ ID NO: 1-911, or 1823-2478, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other  
10 nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a  
15 polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or  
20 eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-911, or 1823-2478 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral  
25 vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-911, or 1823-2478 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the  
30 art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example: Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene), pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); Eukaryotic:

pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al.,  
5 *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means  
10 that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two  
15 appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of  
20 replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among  
25 others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or  
30 simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic

selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., Nat. Biotech 17, 870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intra-muscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

#### 4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1-911, or 1823-2478, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a



sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID NO: 1-911, or 1823-2478 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1-911, or 1823-2478 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences that flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1-911, or 1823-2478, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of an mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of an mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of an mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine,

1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\alpha$ -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15:

6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

#### 5           4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in  
10 Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of an mRNA. A ribozyme having specificity for a nucleic acid of the invention can be designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1-911, or 1823-2478). For example, a derivative of Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the  
15 active site is complementary to the nucleotide sequence to be cleaved in a mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, mRNA of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

20           Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

25           In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids"  
30 or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The

synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.*

86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*,  
5 Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

#### 4.5 HOSTS

10 The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are  
15 in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (*e.g.*, by homologous recombination) to provide increased polypeptide expression by replacing, in  
20 whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also  
25 contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification  
30 of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by

calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to  
5 produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular  
10 polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and  
15 eukaryotic hosts are described by Sambrook, et al., in *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines  
20 of monkey kidney fibroblasts, described by Gluzman, *Cell* 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants,  
25 HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice,  
30 and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used,

as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

5           Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial  
10       strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

15           In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory  
20       sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, and regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein  
25       produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

30           The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory

element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 912-1822, or 2479-3134 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO: 1-911, or 1823-2478 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO: 1-911, or 1823-2478 or (b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 912-1822, or 2479-3134 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as



SEQ ID NO: 912-1822, or 2479-3134 or the corresponding full length or mature protein; and “substantial equivalents” thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 912-1822, or 2479-3134.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., *Bio/Technology* 10, 773-778 (1992) and in R. S. McDowell, et al., *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. Fragments are also identified in Tables 3A, 3B, 5, or 6.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The predicted signal sequence is set forth in Table 6. The mature form of such protein may be obtained and confirmed by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell and sequencing of the cleaved product. One of skill in the art will recognize that the actual cleavage site may be different than that predicted in Table 6. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed (See, e.g., Sakal et al., *Prep. Biochem. Biotechnol.* (2000), 30(2), pp. 107-23, incorporated herein by reference).

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or

from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in *Molecular Cloning: A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO: 912-1822, or 2479-3134.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of

interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™;

one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

#### 4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., *Nucleic Acids Research* 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., *J. Molec. Biol.* 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., *Nucleic Acids Res.* vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., *J. Comp. Biol.*, Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, *ISMB-97*, Vol. 4, pp. 202-209, herein incorporated by reference), Pfam software (Sonnhammer et al., *Nucleic Acids Res.*, Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobicity prediction algorithm (*J. Mol Biol*, 157, pp. 105-31 (1982), the GeneAtlas software (Molecular Simulations Inc. (MSI), San Diego, CA) (Sanchez and Sali (1998) *Proc. Natl. Acad. Sci.*, 95, 13597-13602; Kitson DH et al, (2000) "Remote homology detection using structural modeling – an evaluation" Submitted; Fischer and Eisenberg (1996) *Protein Sci.* 5, 947-955), Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark) incorporated herein by reference).

Polypeptide sequences were examined by a proprietary algorithm, SeqLoc that separates the proteins into three sets of locales: intracellular, membrane, or secreted. This prediction is based upon three characteristics of each polypeptide, including percentage of cysteine residues, Kyte-Doolittle scores for the first 20 amino acids of each protein, and Kyte-Doolittle scores to calculate the longest hydrophobic stretch of the said protein. Values of predicted proteins are compared against the values from a set of 592 proteins of known cellular localization from the Swissprot database (<http://www.expasy.ch/sprot>). Predictions are based upon the maximum likelihood estimation.

Pesence of transmembrane region(s) was detected using the TMPred program ([http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)).

The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215:403-410 (1990).

#### 4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus, or to the middle.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprise one or more domains fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as

appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

#### 4.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.



Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or

stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or

inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model

systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

#### 4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

#### 4.10.1 RESEARCH USES AND UTILITIES

5       The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on  
10       gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel  
15       polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction),  
20       the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

      The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput  
25       screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative  
30       receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

      Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

#### 4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

#### 4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I.  
 5 Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of  
 10 mouse and human interleukin- $\gamma$ , Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current  
 15 Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991;  
 20 Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1,  
 25 John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies,  
 30 E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411,

1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

#### 4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

5 A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

20 It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), 25 neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the



polypeptide of the invention can be used as a feeder layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5 Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotent/pluripotent stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotent/pluripotent mRNA to create  
10 cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present  
15 invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as  
20 well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be  
25 manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into  
30 cardiomyocytes (Wobus et al., *Differentiation*, 48: 173-182, (1991); Klug et al., *J. Clin. Invest.*, 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering* eds. Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence

of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

*In vitro* cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci. U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

#### 4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow

transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are  
5 cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. *Cellular Biology* 15:141-151, 1995; Keller et al., *Molecular and Cellular Biology* 13:473-486, 1993; McClanahan et al., *Blood* 81:2903-2915,  
10 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In *Culture of Hematopoietic Cells*. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994;  
15 Hirayama et al., *Proc. Natl. Acad. Sci. USA* 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In *Culture of Hematopoietic Cells*. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., *Experimental Hematology* 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In *Culture of Hematopoietic Cells*.  
20 R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In *Culture of Hematopoietic Cells*. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc.,  
25 New York, N.Y. 1994.

#### **4.10.6 TISSUE GROWTH ACTIVITY**

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and  
30 tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a

polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced  
5 craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by  
10 blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue  
15 or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or  
20 ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide  
25 environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions  
30 may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central

and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and  
5 localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from  
10 chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

15 Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic  
20 scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

25 A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in:

30 International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

5

#### 4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting  
10 such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from  
15 autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may  
20 be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis,  
25 graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic  
30 contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxicol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self-tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.



Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and  $\beta_2$  microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA

78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk,

Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine  
5 et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

#### 4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related  
10 activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a  
15 contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin  
20 molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

25 The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc.  
30 Natl. Acad. Sci. USA 83:3091-3095, 1986.

#### 4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes.

- 5 Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to
- 10 tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

- A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed
- 15 movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

- Assays for chemotactic activity (which will identify proteins that induce or prevent
- 20 chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margules, E. M. Shevach, W. Strober, Pub. Greene
- 25 Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

#### 30 **4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY**

A polypeptide of the invention may also be involved in hemostasis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders

(including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### 4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness.

Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and

prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Kaposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic

treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

*In vitro* models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wiley-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

#### 4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A. M. Kruisbeek, D. H.

Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions

5 7.28.1- 7.28.22), Takai et al., *Proc. Natl. Acad. Sci. USA* 84:6864-6868, 1987; Bierer et al., *J. Exp. Med.* 168:1145-1156, 1988; Rosenstein et al., *J. Exp. Med.* 169:149-160 1989; Stoltenborg et al., *J. Immunol. Methods* 175:59-68, 1994; Stitt et al., *Cell* 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be

10 identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to

15 radioisotopes, colorimetric molecules or a toxin molecules by conventional methods.

("Guide to Protein Purification" Murray P. Deutscher (ed) *Methods in Enzymology* Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not

limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric

20 molecules. Examples of toxins include, but are not limited, to ricin.

#### 4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening

25 techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays.

30 Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.



Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

5 Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for  
10 screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides  
15 or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see  
20 Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol.*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol*, 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein  
25 permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

30 The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of

the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

#### 4.10.14 ASSAY FOR RECEPTOR ACTIVITY

5           The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another  
10       example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the  
15       invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The responses of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of  
20       the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

25           The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the  
30       chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

#### 4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity.

The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflammation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic myelogenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

#### 4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

#### 4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;
- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human  
5 immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which  
10 elicit any of the following effects may be useful according to the invention:

- (i) increased survival time of neurons in culture;
- (ii) increased sprouting of neurons in culture or *in vivo*;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*,  
*e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or  
15 (iv) decreased symptoms of neuron dysfunction *in vivo*.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or  
20 Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

25 In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited  
30 to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motor-sensory Neuropathy (Charcot-Marie-Tooth Disease).

#### 4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

#### 4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate  
5 fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or  
10 more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect  
15 the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein,  
20 e.g., by an antibody specific to the variant sequence.

#### **4.10.20 ARTHRITIS AND INFLAMMATION**

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The  
25 experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et al., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected  
30 at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed *Mycobacterium tuberculosis* in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of *Mycobacterium* CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

#### 4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

##### 4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1 µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.



#### 4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result,

pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently  
5 administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site).

Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound

10 sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined  
15 amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active  
20 ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with  
25 the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

#### 4.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular,

subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

#### 4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or

other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such

carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, 5 after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, 10 disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. 15 Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler 20 such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the 25 compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, 30 dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be

formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain  
5 formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain  
10 substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in  
15 powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may  
20 also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD  
25 co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics.  
30

Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be

supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

5           The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. 10           Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

          The amount of protein or other active ingredient of the present invention in the 15           pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active 20           ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01  $\mu$ g to about 100 mg 25           (preferably about 0.1  $\mu$ g to about 10 mg, more preferably about 0.1  $\mu$ g to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the 30           therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically



useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition  
5 would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

10 The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides.  
15 Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above-mentioned types of material, such as polylactic acid and  
20 hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize  
25 a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose,  
30 hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful

herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

#### 4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate *in vitro* assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the  $IC_{50}$  as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the  $LD_{50}$  (the dose lethal to 50% of the population) and the  $ED_{50}$  (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between  $LD_{50}$  and  $ED_{50}$ . Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the  $ED_{50}$  with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics

and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 µg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

#### 4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

#### 4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen-binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F<sub>ab</sub>, F<sub>ab'</sub> and F<sub>(ab')<sub>2</sub></sub> fragments, and an F<sub>ab</sub> expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ

from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG<sub>1</sub>, IgG<sub>2</sub>, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

- 5           An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as
- 10 immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence shown in SEQ ID NO: 912-1822, or 2479-3134, or Tables 3A, 3B, 5, or 6, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope.
- 15 Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

- In certain embodiments of the invention, at least one epitope encompassed by the
- 20 antigenic peptide is a surface region of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be
- 25 generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or
- 30 derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind polypeptides of the invention exclusively (*i.e.*, able to distinguish the polypeptide of the invention from other similar polypeptides despite sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow et al. (Eds), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, full-length polypeptides of the invention. As with antibodies that are specific for full length polypeptides of the invention, antibodies of the invention that recognize fragments are those which can distinguish polypeptides from the same family of polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins.

Antibodies of the invention are useful for, for example, therapeutic purposes (by modulating activity of a polypeptide of the invention), diagnostic purposes to detect or quantitate a polypeptide of the invention, as well as purification of a polypeptide of the invention. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific. The invention further provides a hybridoma that produces an antibody according to the invention. Antibodies of the invention are useful for detection and/or purification of the polypeptides of the invention.

Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues in which a fragment of the polypeptide of interest is expressed. The antibodies may also be used directly in therapies or other diagnostics. The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and Sepharose®, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D.M. et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); Jacoby, W.D. et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as for immuno-affinity purification of the proteins of the present invention.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

#### 4.13.1 POLYCLONAL ANTIBODIES

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface-active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as

Bacille Calmette-Guerin and *Corynebacterium parvum*, or similar immunostimulatory agents. Additional examples of adjuvants that can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

5 The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity  
10 chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

#### 4.13.2 MONOCLONAL ANTIBODIES

15 The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus  
20 contain an antigen-binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256, 495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an  
25 immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells  
30 of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, *Monoclonal Antibodies: Principles and Practice*, Academic Press, (1986) pp. 59-



103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused,  
5 immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high  
10 level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for  
15 the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the  
20 binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107, 220 (1980). Preferably,  
25 antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640  
30 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures

such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

#### 4.13.3 HUMANIZED ANTIBODIES

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321, 522-525 (1986); Riechmann et al., Nature, 332, 323-327 (1988); Verhoeven et al., Science, 239, 1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539). In some instances, Fv framework residues of the human

immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or  
5 substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2, 593-596  
10 (1992)).

#### 4.13.4 HUMAN ANTIBODIES

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from  
15 human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Human  
20 monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80, 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques,  
25 including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227, 381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in  
30 humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368, 856-859 (1994)); Morrison (Nature 368, 812-13

(1994)); Fishwild et al, (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13, 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals that are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse<sup>TM</sup> as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells that secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that

contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

5 In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

#### 10 4.13.5 FAB FRAGMENTS AND SINGLE CHAIN ANTIBODIES

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of  $F_{ab}$  expression libraries (see e.g., Huse, et al., 1989 Science 246, 1275-1281) to allow rapid and effective  
15 identification of monoclonal  $F_{ab}$  fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an  $F_{(ab)2}$  fragment produced by pepsin digestion of an antibody molecule;  
(ii) an  $F_{ab}$  fragment generated by reducing the disulfide bridges of an  $F_{(ab)2}$  fragment; (iii) an  
20  $F_{ab}$  fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv)  $F_v$  fragments.

#### 4.13.6 BISPECIFIC ANTIBODIES

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies  
25 that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the  
30 recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305, 537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas)

produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10, 3655-3659.

5       Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of  
10       the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh *et al.*, *Methods in Enzymology*, 121, 210 (1986).

      According to another approach described in WO 96/27011, the interface between a  
15       pair of antibody molecules can be engineered to maximize the percentage of heterodimers that are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or  
20       similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

      Bispecific antibodies can be prepared as full-length antibodies or antibody fragments  
25       (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brenman *et al.*, *Science* 229, 81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent  
30       sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB

derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., *J. Exp. Med.* 175, 217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., *J. Immunol.* 148(5), 1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90, 6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., *J. Immunol.* 152, 5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., *J. Immunol.* 147, 60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen.

Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA.

Another bispecific antibody of interest binds the protein antigen described herein and further  
5 binds tissue factor (TF).

#### 4.13.7 HETEROCONJUGATE ANTIBODIES

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such  
10 antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction  
15 or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

#### 4.13.8 EFFECTOR FUNCTION ENGINEERING

It can be desirable to modify the antibody of the invention with respect to effector  
20 function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-  
25 mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176, 1191-1195 (1992) and Shopes, J. Immunol., 148, 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53, 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can  
30 thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3, 219-230 (1989).

#### 4.13.9 IMMUNOCONJUGATES



The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

5 Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, *Phytolaca americana* proteins  
10 (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include  $^{212}\text{Bi}$ ,  $^{131}\text{I}$ ,  $^{131}\text{In}$ ,  $^{90}\text{Y}$ , and  $^{186}\text{Re}$ .

Conjugates of the antibody and cytotoxic agent are made using a variety of  
15 bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates  
20 (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

25 In another embodiment, the antibody can be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

#### 4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO: 1-911, or 1823-2478 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO: 1-911, or 1823-2478 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes.

Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search

algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein-encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

5           As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently  
10       available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence  
15       information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

          As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural  
20       motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but  
25       is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more  
30       amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized

that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

#### 4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix-see Lee et al., Nucl. Acids Res. 6, 3073 (1979); Cooney et al., Science 15241, 456 (1988); and Dervan et al., Science 251, 1360 (1991)) or to the mRNA itself (antisense-Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

#### 4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the

polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the

invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

#### 4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

#### 4.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a

polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO: 1-911, or 1823-2478, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
- (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be

selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein  
5 encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in  
10 order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly  
15 described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of  
20 DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are  
25 designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6, 3073 (1979); Cooney et al., Science 241, 456 (1988); and Dervan et al., Science 251, 1360 (1991)) or to the mRNA itself (antisense-Okano, J. Neurochem. 56, 560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in  
30 a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of



the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

#### 4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO: 1-911, or 1823-2478. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from any of the nucleotide sequences SEQ ID NO: 1-911, or 1823-2478 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well-known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The

technique of fluorescent *in situ* hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

#### 4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6), 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8), 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridgeheads for further covalent coupling.

CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) *Anal. Biochem.* 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen *et al.*, (1991). In this technology, a phosphoramidate bond is employed (Chu *et al.*, (1983) *Nucleic Acids Res.* 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ $\mu$ l) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm<sub>7</sub>), is then added to a final concentration of 10 mM 1-MeIm<sub>7</sub>. A ss DNA solution is then dispensed into CovaLink NH strips (75  $\mu$ l/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm<sub>7</sub>, is made fresh and 25  $\mu$ l added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described

by Fodor *et al.* (1991) Science 251(4995), 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res., 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1), 104-8; all references being specifically incorporated herein.

5 To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) Proc. Nat'l. Acad. Sci., USA 91(11), 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256  
15 spatially defined oligonucleotide probes may be generated in this manner.

#### 4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook  
20 *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA  
25 samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990)  
30 Nucleic Acids Res. 18(24), 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the

cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*JI, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and  
5 fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *Cvi*JI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the  
10 specificity of this enzyme (*Cvi*JI\*\*), yield a quasi-random distribution of DNA fragments from the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *Cvi*JI\*\* digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *Cvi*JI\*\* restricts  
15 pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 µg instead of  
2-5 µg); and fewer steps are involved (no preligation, end repair, chemical extraction, or  
20 agarose gel electrophoresis and elution are needed).

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are  
25 contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

#### 4.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which  
30 correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm<sup>2</sup>, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns,

separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm<sup>2</sup> and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

## **5.0 EXAMPLES**

### **5.1 EXAMPLE 1**

#### **Novel Nucleic Acid Sequences Obtained From Various Libraries**

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human

chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences.

## 5.2 EXAMPLE 2

### Assemblage of Novel Nucleic Acids

The contigs or nucleic acids of the present invention, designated as SEQ ID NO: 1823-2478 were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST, gb pri, and UniGene, and exons from public domain genomic sequences predicated by GenScan) that belong to this assemblage. The algorithm terminated when there were no additional sequences from the above databases that would extend the assemblage. Further, inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

## 5.3 EXAMPLE 3

### Novel Nucleic Acids

The novel nucleic acids of the present invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The nucleic acids were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (Hyseq's database containing EST sequences, dbEST, gb pri, and UniGene) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences

into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full-length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any  
 5 frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequences were checked using FASTY and/or BLAST against Genbank (i.e., dbEST, gb pri, UniGene, and Genpept) and the Geneseq (Derwent). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and cg-zip-2 (Hyseq, Inc.). The full-length nucleotide and amino acid  
 10 sequences, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NO: 1-1822.

Table 1 shows the various tissue sources of SEQ ID NO: 1-911.

The homologs for polypeptides SEQ ID NO: 912-1822, that correspond to nucleotide sequences SEQ ID NO: 1-911 were obtained by a BLASTP version 2.0a1 19MP-WashU  
 15 searches against Genpept and Geneseq (Derwent) using BLAST algorithm. The results showing homologues for SEQ ID NO: 912-1822 from Genpept 127-129 are shown in Table 2A. The results showing homologues for SEQ ID NO: 912-1822 from Genpept 131 are shown in Table 2B.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J.  
 20 Comp. Biol., Vol. 6, 219-235 (1999), <http://motif.stanford.edu/ematrix-search/> herein incorporated by reference), all the polypeptide sequences were examined to determine whether they had identifiable signature regions. Scoring matrices of the eMatrix software package are derived from the BLOCKS, PRINTS, PFAM, PRODOM, and DOMO databases. Table 3 shows the accession number of the homologous eMatrix signature found  
 25 in the indicated polypeptide sequence, its description, and the results obtained which include accession number subtype; raw score; p-value; and the position of signature in amino acid sequence. The results showing homologous signatures for SEQ ID NO: 912-1822 from eMatrix version 1.0 are shown in Table 3A. The results showing homologous signatures for SEQ ID NO: 912-1822 from eMatrix version 2.0 are shown in Table 3B.

30 Using the Pfam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4A shows the name of the Pfam model found, the description, the e-value and the Pfam score for the



identified model within the sequence using Pfam versions 7.0 and 7.2. Table 4B shows the name of the Pfam model found, the description, the e-value and the Pfam score for the identified model within the sequence using Pfam version 7.5. Further description of the Pfam models can be found at <http://pfam.wustl.edu/>.

5           The GeneAtlas™ software package (Molecular Simulations Inc. (MSI), San Diego, CA) was used to predict the three-dimensional structure models for the polypeptides encoded by SEQ ID NO: 1-911 (i.e. SEQ ID NO: 912-1822). Models were generated by (1) PSI-BLAST which is a multiple alignment sequence profile-based searching developed by Altschul et al, (Nucl. Acids. Res. 25, 3389-3408 (1997)), (2) High Throughput Modeling  
10 (HTM) (Molecular Simulations Inc. (MSI) San Diego, CA,) which is an automated sequence and structure searching procedure (<http://www.msi.com/>), and (3) SeqFold™ which is a fold recognition method described by Fischer and Eisenberg (J. Mol. Biol. 209, 779-791 (1998)). This analysis was carried out, in part, by comparing the polypeptides of the invention with the known NMR (nuclear magnetic resonance) and x-ray crystal three-dimensional structures  
15 as templates. Table 5 shows: "PDB ID", the Protein DataBase (PDB) identifier given to template structure; "Chain ID", identifier of the subcomponent of the PDB template structure; "Compound Information", information of the PDB template structure and/or its subcomponents; "PDB Function Annotation" gives function of the PDB template as annotated by the PDB files (<http://www.rcsb.org/PDB/>); start and end amino acid position of  
20 the protein sequence aligned; PSI-BLAST score, the verify score, the SeqFold score, and the Potential(s) of Mean Force (PMF). The verify score is produced by GeneAtlas™ software (MSI), is based on Dr. Eisenberg's Profile-3D threading program developed in Dr. David Eisenberg's laboratory (US patent no. 5,436,850 and Luthy, Bowie, and Eisenberg, Nature, 356:83-85 (1992)) and a publication by R. Sanchez and A. Sali, Proc. Natl. Acad. Sci. USA,  
25 95:13597-12502. The verify score produced by GeneAtlas normalizes the verify score for proteins with different lengths so that a unified cutoff can be used to select good models as follows: Verify score (normalized) = (raw score – 1/2 high score)/(1/2 high score)

          The PFM score, produced by GeneAtlas™ software (MSI), is a composite scoring function that depends in part on the compactness of the model, sequence identity in the  
30 alignment used to build the model, pairwise and surface mean force potentials (MFP). As given in table 5, a verify score between 0 to 1.0, with 1 being the best, represents a good model. Similarly, a PMF score between 0 to 1.0, with 1 being the best, represents a good model. A SeqFold™ score of more than 50 is considered significant. A good model may

also be determined by one of skill in the art based all the information in Table 5 taken in totality.

Table 6 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et al reference, was obtained for the polypeptide sequences.

Table 7 correlates nucleotide sequences of the invention to a specific chromosomal location when assignable.

Table 8 shows the number of transmembrane regions, their location(s), and TMPred score obtained, for each of the SEQ ID NO: 912-1822 that had a TMPred score of 500 or greater, using the TMpred program ([http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)).

Table 9 is a correlation table of the novel polynucleotide sequences SEQ ID NO: 1-911, their corresponding polypeptide sequences SEQ ID NO: 912-1822, their corresponding priority contig nucleotide sequences SEQ ID NO: 1823-2478, their corresponding priority contig polypeptide sequences SEQ ID NO: 2479-3134, and the US serial number of the priority application (all of which are herein incorporated in their entirety), in which the contig sequence was filed.

Table 10 is a correlation table of the novel polynucleotide sequences SEQ ID NO: 1-911, the novel polypeptide sequences SEQ ID NO: 912-1822, and the US application serial number and corresponding SEQ ID NO in which the sequence was previously filed.

Table 1  
115

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
*Mixture of 16 tissues - mRNA	Various Vendors	CGd011	58 74 147 267 277 432 489 496 618
*Mixture of 16 tissues - mRNA	Various Vendors	CGd012	3-4 6 9 19-20 41 56-58 62 71-72 74 78 81 91-92 114 118-119 128 133 141 147-149 153 155 165-167 176-177 185 187 195 215 218 220 228-229 237 245 249 257 259 267 271 277-280 302 310 313 326 353 355 366 368-369 380 385 391 394 398 405-406 409-410 417 432 439-440 442 449 452 460 462 464 471-472 475 480 489 496 498 505 524- 528 535 546 555 581 585 594 610 618 636 804 811 816 845 871
*Mixture of 16 tissues - mRNA	Various Vendors	CGd013	55 71 78 92 165 228 271 276 279-280 310 353 355 422 424 496 526 531 606 618 908
*Mixture of 16 tissues - mRNA	Various Vendors	CGd015	10 97 100 194-195 246 258-260 266 270 282 311 355-357 398 408 423 428- 430 439 495 526 544 805
*Mixture of 16 tissues - mRNA	Various Vendors	CGd016	10 14 20 51 71 75 83 116 139 149 168 174-175 194-195 209 230 233 238-239 248 257 266-267 279-280 289 311-312 355 403 420-421 493 503 513 515 532 543 689 707 909
*Mixture of 16 tissues - mRNA	Various Vendors	CGd010	176 185 352 380 386 403 422 425-426 517 537 606
adrenal gland	Clontech	ADR002	7 21 36 40 43 64 69 88 101 120 130 147 157 162-163 169 184 187 195 203 218 225 230 236 275 283 289 298-299 317 355 381-382 392 425 427 430 446 449 460 471 477 503 510 524 531 536 540 544 578 622 627 648 668 729 788-789
adult bladder	Invitrogen	BLD001	19 33 54 128 195 220 312 355 395 421 423 446 464 503 543 618 829
adult brain	BioChain	ABR012	184 246 355 364
adult brain	BioChain	ABR013	102 226 355 423 432
adult brain	GIBCO	AB3001	3 87 97 100 236 269 304-306 355 377 427 497 536 540
adult brain	GIBCO	ABD003	16 20-21 27 40 59 62 87 94 100 105 117 120 128-129 134 140 157 184 191- 192 195 201 234 238-240 245-246 250 256 260 267-268 270 272 295 304-306 316 335 351-353 355 362 369 377 384 399 412 415 419 423 430 440 457-460 464 470 477 496 498 503 519 536 539- 540 543 555 589 637
adult brain	Invitrogen	ABR014	120 199 226 246 248 312 319 355 401 445 477 503 611
adult brain	Invitrogen	ABR015	20 27 42 49 100 135 157 186 195 199 226 246 256 430 432 446 472 484 543 868
adult brain	Invitrogen	ABR016	16 100 120 195 226 267 355 459 496 526
adult brain	Clontech	ABR001	21 27 40 59 87 101 112 128 226 268

Table 1  
116

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			272 355 364 382 387 428-429 464 503 517 543 619-620 644 838 847 852
adult brain	Clontech	ABR006	17 23 30 33 41 43 59 64 88 94 98 120- 121 124 131 133 154 179 183-184 194 199 207 214 220 226 231 238-240 250 256 257 266 278 281 304 312 351 353 355 365 376 382 392 401 412 415 419 427 430 433 445 460 464 479 516-517 519 535-537 540 542-543 546 548 555 582 728 747-748 754 816
adult brain	Clontech	ABR008	3 17 23 26 29 31 36 40-41 44 49 51 57 59 62 78 83 87-88 90 100-102 109 115 121 124 127-128 130 131 133-136 145- 146 149 154 158 161-163 167 169-171 182 184 186 195-196 202 207 211 214 224 227 231 244-246 248 250-251 256 261-262 267 276 281 285-286 296-297 312 319 337 351 353 355 364-365 368 372 376 380 383 385 390-391 394 401 403 412 423 427 431 433 435-436 437 444-445 447-448 452 457-458 460-461 465-466 479-481 489 503 506 508 513 517 526 535 545 577-578 585 592 598 611 621 624-626 648 674-675 722 816 833 837 839 847-848 852 857 865-866 884
adult brain	Clontech	ABR011	26 285-286 431
adult brain	Invitrogen	ABT004	19-21 30 35 44 57 59 64 87 128 131 140 144 157 161 187 198 202 226 230 250 268 272 293 295 351 355 365 372 387 395 415 424 427 428 442-443 446 452 455 457-458 478 498 513 516 524 540 543 816 831
adult cervix	BioChain	CVX001	10 22-23 36 41 43 62-63 77 100-101 105-106 109 120 130 134 137 141 154 160 162-163 178-180 183-184 187 192 195 198 215 217 226 240 246 249 260 262 267 272 285 297 304 310-311 316 334-336 353 355 373-374 377 379-380 385 395 400 408-409 412 417 425 431 452 464 468 476-477 482 496-497 503 508 512 516 540 546 578 592 626 645 660-662 676 680 888
adult colon	Invitrogen	CLN001	19 21 57 109 128 130 131 148 168 230 240 301 364 412 424 440 452 464 469 543 664
adult heart	GIBCO	AHR001	3 19-20 22 26 41 45 59 62 67 94 97 99- 101 106 109 120 124 126 128 134 138 157 171 183 186-190 195 197-198 202 205 211 215 226-227 238-240 248 249 261-262 267-268 281 285 311 316 335 353 373-374 379-380 399 412 419 423 425 430 433 440 464 477 480 496 498 503 506 543 546 553 566 570 580
adult kidney	GIBCO	AKD001	3 17 19-20 25 32 35-37 41 49 57 59 62

Table 1  
117

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			64 87 94 98-102 105 109 120 128 134-137 145 153-154 157-158 160 176 179-180 182 184 186 192 195 198 215 226-227 234 238-240 246 248 256 257 262 267 272 278 283 285 297 304 310-311 316 326 353 355 362 364 372 375 377 379 381-382 391 395 412 415 421-422 423-424 430-432 434 437 440 444 452 459 464 470 477 496 503 506 510 519 527-528 540 546 551 566 578 585 606 837
adult kidney	Invitrogen	AKT002	1 17 20 26 59 62 73 87 94 100-101 120 133 135 154 160 177 187 195 208 226 228 238-239 248 249 257 260 262-264 267-268 276 293 299 353 355 359 362 380 412 419 421 425 427 431 434 436 439 452 462 464 498 503 508 515 519 532 540 543 546 586 590
adult liver	Invitrogen	ALV002	1 3-4 14-16 19-20 32-33 40 48 54 63 77 87 94 97 101 128-129 135-136 138-139 149 157 180 182 187 193 195-196 203 226-227 246 248 257 262 277-278 305-306 316 347 355-357 362 364 383 391 412 423 427 430 446 477 496 503 510 525 540 553 629 636 852
adult liver	Clontech	ALV003	14 24 34 40 94 160 195 227 257 278 355-357 362 369 424 436 496 527-528 750 901
adult lung	GIBCO	ALG001	20 36 40 100 105 120 128 130 154 157 195 198 263-264 267 273 336 373-374 420 437 446 464 477 503 543
adult ovary	Invitrogen	AOV001	1 13 15 17 19-22 36 41 49 51 57 59 62 68 70-71 77 87-88 97 100-101 105-106 109 112 116 120 125 128 130 134 144 157-160 169 176-177 179 181-182 184 187 191-193 195 198 203 211 213 215 226 230 236-237 240 246 248 249 261-264 267-268 271-272 279-280 288-289 295 298-299 304 310 320 335 355 362 364-365 376-377 380 383-386 389 391 394-395 399 408 412 415 417 422 424 434 436 440 446 452 456 459-460 477 479 496-497 503 506 510 518-519 535-536 540 543-544 546 555 591 612-614 794 823 827 829-830 833-834 860
adult placenta	Clontech	APL001	16 195 256 285 399 421 424 446 464 477 792
adult spleen	Clontech	SPLc01	17 22 36 46 88 91 93 97 103 112 135-136 142 147 162-163 187 195 256 263-264 268 281 303-306 311-312 353 376 427 433 436 437 452 477 480 506 515 518 535 537 548
adult spleen	GIBCO	ASP001	3 10 46 77 81 94 97 100 105 120 129 134-136 160 166 186-187 195 226 240

Table 1  
118

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			246 248 256 257 267 303-304 311 316 355 377 382 389 391 423 427 430 432 436 452 477 536 540 543 582 586
adult testis	GIBCO	ATS001	20-21 100-101 105-106 109 128 154 157 160 169 191-193 195 198 213 248 256 267 269 295 316 336 355 380 388 415 452 460 477 496 508 518 540-541
bone marrow	GF	BMD002	10 17 20 25 36 41 43 45-46 58-59 70- 71 75 81 86 88 100 106 109 115 120 124 130 134 136 138 147 154-155 157 162-163 166-167 182-184 187-190 193 195-196 198 202 216-217 220 226 231 233 240 248 256 258-259 267-268 272 279-280 338-339 359 364 373-374 376-377 380 382 389 391 398 403 412 415 417 423 427 432 436 437 444 446 452 460 464 477 480 490 493 497 503 505 513 517-518 526-528 532 537 540 543 546 548 551 581-582 597 613 626 694 697 704 723 729 741-742 744 751 754-755 768 775 815 822 855 859 874 885 891
bone marrow	Clontech	BMD001	10 16 20 22 25 36 41 49 59 62 71 81 97 99-100 120 128 130 134 136 143 157 179 184 187-190 195 198 207 211 215- 217 236 247 256 258-260 263-265 267 269 279-280 282 285 298 304 312 353 364 373-375 377-378 380 389 398 419 421-422 423 425 430 432 434 452 477 480 497 503 506 518-519 536-537 544 546 551 556 566 569 571-572 602 604 615 728 815 818
bone marrow	Clontech	BMD004	259 398
bone marrow	Clontech	BMD007	36 259 398
bone marrow	BMD008	259	398
bone marrow	null	STM001	91 136 166 195 256 362 377 412 423- 425 477 503 537 540
CGSP009		255	
cultured preadipocytes	Stratagene	ADP001	9 17 76 100 120 134 136 179 195 226 246 248 263-264 267 275 281 295 323 345-346 377 391 423-425 427 430-431 464 477 503 506 531 543 546 548 592 701 876
endothelial cells	Stratagene	EDT001	3-4 10 20 22-23 35-36 41 47 51 57 59 62 64 87 97 99-102 105-106 108-110 120 128 134-135 144 153 157 167 177 179 182-183 186 191 195 204 214-215 226 240 246 248 251 256 262 267-268 270 272-273 275 282 285 295 304 353 355 362 364-365 373-374 376-377 379 382 385-386 388 391 394 403 412 419 423-425 427 430-432 440 444-445 477 497 503 506 508 546 548 566 608
esophagus	BioChain	ESO002	532 607
fetal brain	Invitrogen	FBT002	3 20 42 57 64 100 130 134-135 154

Table 1  
119

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			157 162-163 169 214 226 250 262 265 267 272 283 294 336 351 355 365 407 412 423 431 440 444-445 460 470 477 498-499 526 544 578 610
fetal brain	GIBCO	HFB001	4 20 22 26-31 36 42 49 59 64 77 85 87 97 100-102 105 108-109 120 128 134 157 179 184 186 191-192 195 198-199 211 226 238-239 246 248 256 261-262 267-268 285 295 300 310 335 353 355 365 376-377 379-380 390 415 419 421 423-424 430 446 448 460 477-478 489 496 503 506 513 516 536 542 546 595 793 817
fetal brain	Clontech	FBRs03	49
fetal brain	Clontech	FBR001	30 87 100 182 199 268 415 435 516 578
fetal brain	Clontech	FBR004	36 214 224 229 376 394 445 460 513 516 578 611 787 816 861 897
fetal brain	Clontech	FBR006	3 12 15 20-21 23 25 33-34 36 44 49 58- 59 62 78 83 88 91 99-100 103 106 124 128 130 131 134-136 149 154 161 169 182 184 186-187 193 227 231 236 242 245 256 261 281 285 292 301 311-312 320 353 355 360-361 364-365 376 383 385 401-403 412 422 423-424 427 433 435 441 444 446 452 457-458 460 465 471 478-480 486 496 502-503 506 513 515 517 526 535 537 546 548 566 756- 757 759-760 784-785 858-859 886
fetal heart	Invitrogen	FHR001	3 7 12 16 19-20 26 31 84 94 101-102 106 109 124 135-136 138 142 158 183- 184 193 195 227 229 236 241 249 265 267 270 273 277 285 307 310 353 355 376 380 382 385 387 394 396 408 412 419 425-426 430-432 444-445 452 464 477 498 503 506-509 513 526 543 546 548 592 598 648 761-763 765-766 776-778 859-860 893
fetal kidney	Invitrogen	FKD007	26 111 162-163
fetal kidney	Clontech	FKD001	46 186-187 195 267 335 421 452
fetal kidney	Clontech	FKD002	12 44 49 71 93 97 100 102 106 128 130 136 154 162-163 179 183 186 195 227 267 276 279-280 285 310 316 376 384 387 389 464 477 480 496 498-499 510 535 554 582 729 768 894 897
fetal liver	Clontech	FLV002	20 34 40 68 78 85 93 125 182 194-195 207 257 259 266 278 297 356-357 383 398 457-458 477 496 548 767
fetal liver	Clontech	FLV004	5 12 14 20 24 32-33 40 49 94 100 115 128 130 135 157 159-160 169 176 179- 180 195-196 203 224 226-227 229 246 249 253 257 259 267 278 295 305-306 355-357 385 398 424 427 437 459 464 477 496 498 537 546 581 603 641 743 768

Table 1  
120

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
fetal liver	Invitrogen	FLV001	20 32-33 36 85 94 128 144 149 154 157 171 180 182 195 248 278 292 312 316 355 362 423 430-431 436 441 472 477 498 540 543 555
fetal liver-spleen	Columbia University	FLS001	2-3 10 12 14-17 19-24 26-27 32-36 39 41 43 49 51 57 59 62 80 90-94 96-97 99-100 105 107 109 120 128-130 134 136 149 153 157-158 162-163 167 176- 177 179-187 191 194-195 198 211 215 226-228 230 234 237 240 243 245-246 248 254-259 262 266-268 270 277- 278 282-283 285 295 297 310-311 316 320-326 330-331 335-336 338 353 355-357 362 370-374 376 380 382 384 391 394 398 412-413 415 419-421 423 430 439-440 446 452 456 459-460 464 472 477 479-480 489 496-499 503 506 510 515 525-526 534-535 538 540 543- 544 546 548 555 557-560 562-566 568 573-574 576 578 591 593 600-601 732- 735 809-812 814 820 827
fetal liver-spleen	Columbia University	FLS002	2-3 6 10 14-15 17-20 22-24 26 33 37 41 48-49 56-57 62 69 77 81 85-86 90 92 94 99-100 104 107-109 114 120 129 134-135 144 149 157 168-169 174-175 179-181 185-187 194 211 215 224 227 230-231 233 237 240 245-246 257-259 262 266-268 270 272 275 278 298 300 304-306 311 313 316-317 320 326-336 352-354 362 364-365 376-377 382-384 398 411-412 415 419 421 425 430 433 436 439 441 446 452 456 459-460 464 477 479 491-492 496 499 503 505-506 508 510 515 519 526 535 539 544 546 566 575 578 591 593 601 654-657 659 685 687-691 705 707 709 774 779-780 809 812 820 833 837 854-856 872 896
fetal liver-spleen	Columbia University	FLS003	14 17 20 26 34 63 86 99 107 167 179 187 195 211 220 227 231 248 257 275 278 320-321 328 335 355-357 365 384 394 411-412 416 423 425-427 430-431 440 464 486 497 506 510 525 546 558 566 648 732 735 738-740 746 779-781 787 809 812 814 837
fetal lung	Clontech	FLG001	22 49 160 176 195 237 259 267 353 398 421 430 436 456 815
fetal lung	Clontech	FLG004	12 41 506
fetal lung	Invitrogen	FLG003	15 64 115 128 135 142 151 154 186- 187 193 195 220 236 253 276 355 421 446 464 506 519 527-528 537 578 626
fetal muscle	Invitrogen	FMS001	20 94 100 128 138 171 184 195 226 246 259 267 311 398 412 443 478 498 531 585 610
fetal muscle	Invitrogen	FMS002	12 21 38 57 71 86 100-101 104 109 116 129-130 138 179 195-196 204 253



Table 1  
121

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			285 312-313 336 376 384-385 391 403 423 427 430 446 452 473 498 507 517 535 546 598 643 892
fetal skin	Invitrogen	FSK001	2 6 12 30-31 56 59 62-63 74 88 97 100 109 115-118 128 134-135 138 141 151- 152 154 160 168 178 181-182 185 187 193 195 229 236 248 249 252 267 272- 273 276 282 301 313-316 320 362 364- 365 385 391 404-407 410 415 421 423 427 430 446 452 460 471 473-474 477 486-489 499 503 512 526-528 537 540 543 548 551-552 594 598 635 681-682 813 829 870-871 888
fetal skin	Invitrogen	FSK002	9 12 17 20 33 46 58 74 76-78 83 103 118 124 128-129 135 142 162-163 167- 169 177 193 195 203 229 231 244 249 254 256 276 311 313-314 342 346 363 372 376 382 391 393-394 410 421 423 427 433 464 473 477-478 480-481 486- 487 497-499 511-513 515 518 535 537- 538 543 589 592 626 701 731 768-769 771 781 895 898-899
fetal spleen	BioChain	FSP001	46 256
induced neuron-cells	Stratagene	NTD001	15 20 49 87 109 119-120 157 179 195 208 226 234 238-239 267 282 340-344 364 494-495 504 506 516 546 597 698- 699 887
infant brain	Columbia University	IB2003	21-22 26 31 48 77 82 98-99 103 128 131 134-135 154 179 198 211 214 230 234 295 312 339 348-349 351 353 355 362 365 391 415 423 431 445 452 460 470 478 497 513 515-516 535 537 540 543 566 578 639-640 726 816 833 849
infant brain	Columbia University	IBS001	21 59 77 98 131 154 224 244 371 513 540 543 578 639 812
infant brain	Columbia University	IBM002	19 26 77 248 516
infant brain	Columbia University	IB2002	19 21-22 26-27 31 36 57 62 77 93 97- 98 120 128-129 131 134 140 149 154 169 179 182 187 208 214 226 230 234 237 256 267-268 273 281 284 295 336 351 353 355 365 371 387 415 421 429 431 435 445-446 452 460-461 478-479 496 503 508 513 515 536-537 543 546 555 578 605 613 639-640 816 826 849 853
leukocyte	GIBCO	LUC001	3 20 23 35-36 39 59 62 64 70-71 87-88 94 100-101 105-106 120 128-130 134- 137 139 148 154-155 160 165 176 179 184 186-190 192-193 195-196 198 211 216-217 225-226 236 240 246 251 256 257 260 262-264 267-268 270 279-280 285 298 304 311 353 355 362 375 377 380 412 415 417 423-426 432 437 440 444 452 456 460 477 479 497 503 505-506 508 519 540 543 546 565 582-

Table 1  
122

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			585 587-588 596 646 821
leukocyte	Clontech	LUC003	41 47 59 71 100 144 186-187 195 240 263-264 267 279-280 298 385 432 437 446 497 503
lung tumor	Invitrogen	LGT002	20 35-37 41 43-44 47-49 54-55 59 62 64 71 77 93-94 100 102 106 120 134 136 139 154 157 176 186 194-195 198 211-212 215 218 222 226 234 237-239 246 249 256 257 260 262-264 267 276 279-280 289 305-306 311 336 355 362 364 380 382 385 388 394 396 417 419 421 425 427 432-433 438 444 446 449 477 483-485 496-497 503 506 508 527- 528 540 543 589 606 611 616-617 651 679 799 828 835 867-869
lung, fibroblast	Stratagene	LFB001	3-4 16 22 35-36 41 94 100 105 116 120 128 134-135 177 191 195 235 256 267 353 355 364 377 379 425 459 464 477 497 503 703 813 829
lymph node	Clontech	ALN001	10 53 57 81 136 182 195 198 209 246 263 267 353 355 389 432 503 506
lymphocyte	null	DGD001	267 311 373-374 412 430-432 460 464 480 503 536
lymphocytes	ATCC	LPC001	35-36 41 57 77 99-100 105 115 120 128 139 148 155 171 176 179 182 187- 190 192 196 226 232-233 236-237 260 263-264 267-268 270 379 417 427 450 452 497-498 514 518 540 547-548 566 583 587 619 693 851
macrophage	Invitrogen	HMP001	87 128 177 187 196-197 257 277 355 362 385 412 415 420 423 446 464 497 537 543 551 782-783 795
mammary gland	Invitrogen	MMG001	8 16 20 36 44 56-57 59-60 64 74 77 90 94 96-97 100 109 120 128 130 134-136 154 157 160 162-163 168 179 182 187 193 195 198 211 219-220 226 240 248 249 251 257 262 267-268 272-273 276 281 287 300-301 316 355 362 364 372 383 388 393-394 417 421 423 427 430- 432 440 446 452 460 477-478 497 503 506 516 527-528 537 540 543 555 585 597 680 842-843
melanoma from-cell-line- ATCC-#CRL Clontech	MEL004	3 20 26 41 59 100 120 130 160 195 223 237 246 267 269 275 305-306 376 409 412 419 423 437 446 459 464 477	

Table 1  
123

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
		486 647 669 862	
neuronal cells	Stratagene	NTU001	21 49 51 57 59 83 96 120 134 179 195 230-231 256 262 267 289 351 355 383 412 423 427 497 503 527-528 727 875
pituitary gland	Clontech	PIT004	16 19 68 87 100 112 128 192 195 246 262 268 275 355 372 394 416 427 430- 431 477 503 578
placenta	Invitrogen	APL002	16 36 57 77 141 187 230 248 262 295 355 423 430-431 446 496 527-528 555
placenta	Clontech	PLA003	20 33 70 88 90 109 115 120 124 129 141 181 193 237 245 249 285 297 316 355 382 391 403 411 421 446 464 477 486 497 513 548 772-773
prostate	Clontech	PRT001	10 20 36 49 85 94 100 135 142 188- 191 195 198 237 267 282 285 316 320 353 379 394 421 423 425 437 460 506 589 631-632
rectum	Invitrogen	REC001	35 77 120 128 136 153 159 168 182 256 268 272 301 423 430 494 499 503 505 537 543-544 552 578 664
retinoic acid- induced- neuronal-cells	Stratagene	NTR001	120 130 179 183 193 256 355 460 478 503 890
saliva gland	Clontech	SALs03	475
salivary gland	Clontech	SAL001	19 38 66 134-136 139 191 195 206 240 246 256 311 355 362 376 382-383 395 421 432 456 460 464 475 477 494 497-498 503 527-528 666 790 826
skeletal muscle	Clontech	SKM001	17 41 100 195 197-198 211 215 236 238-239 298 302 369 415 443 452 475 496 503
skin fibroblast	ATCC	SFB001	49 464
skin fibroblast	ATCC	SFB002	49 246 267 477
skin fibroblast	ATCC	SFB003	49
small intestine	Clontech	SIN001	3 6 10 46 57 59 62 77 90 94 97 100 102 106 109 112 115 120 123 128 142 162- 163 168 177 179 183 186 188-190 195 224 226 240 246 248 262 267-268 275 289 302 311 316 324 353 355 364 372- 376 380 382-384 394 408 412 423 425 430 432 439 442 444 452 459 464 477- 478 499 503 532 536 540 547-548 626- 627 633 648 729 736-737 745-746 823 828 844
spinal cord	Clontech	SPC001	3 17 27 34 41 57 64 87 100-101 128 157 183 187 192 195 199 204 226 229 234 240 246 248 260-261 263-264 267 304 308-309 311 355 369 382 384 388 412 415 419 428-429 434 446 457-460 478 496 503 519 648 652
stomach	Clontech	STO001	36 57 104 115 120 134-135 168 195 267 311 330-331 364 414 423 425 452 456 460 540 886
thalamus	Clontech	THA002	19 87 113 121 160 182 195 202 248

Table 1  
124

Tissue Origin	Library/RNA Source	HYSEQ Library Name	SEQ ID NOS:
			249 262 272 283 289 355 365 424 427 451 457-458 460 477 480 503 535-536 543 634 841
thymus	Clontech	THMc02	3 12 20 36 58-60 70-72 90 93 97 101- 102 109 114-115 130 135 147-148 160 162-163 167 169 171 179 183-184 187 195-196 217 237 249 259 276 279-280 303 307 312 320 355 376 383-384 398 402 419 422 432-433 436 437 444 464 471 479-480 506 508 513 515 535 537 550 589 641-642 671-672 787 830 858- 859 864
thymus	Clontech	THM001	20 53 59 71 100-101 105-106 110 114 128 130 134 154 157 179 188-191 195 215 224 230 240 248 259 267 279 335 380 382 385 391 398 408 412 427 432 437 452 459 476 497 543 546 611 833 900
thyroid gland	Clontech	THR001	3 16 19-20 23 33 35 41 43-44 48-49 57 59 87 94 100-101 120 133-135 147 149 157 162-163 179 187 191 193 195 198 206 209 215 217 229 246-247 256 261- 262 267 272 276 283 287 297 304-306 309 312 316 335 353 355 362 364 371- 374 380 385 394 399 412 419 421 431 442 444 452-453 460 462 475 477 479 496-497 503 506 519-520 536 543 546 548 596 649 653 704 731 812 845 851
trachea	Clontech	TRC001	100-101 120 130 184 195 267 300 355 380 396-397 399 415 421 423 464 475 503 527-528 546 709 725
umbilical cord	BioChain	FUC001	7 26 36 40 43 46 50-51 57 68 71 77-78 96 100 102 109 115 120 128 142 158- 159 172-173 179 184 186 191 195 198 215 226 228 230 236-237 240 246 250 267 273 275 279-280 283 287 300 304 335 351 355 364-365 381 388 394 399 408 425 430 440 452 459 464 475 477 503 536-537 540 543 546 576 589 630 635 669 682 684 904
uterus	Clontech	UTR001	16 22 64 77 87 100 128 134 195 226 249 335 366 412 425 427 452 454 464 546 551
young liver	GIBCO	ALV001	5 14 20 26 32-33 57 71 87 94 97 100 102 105 120 128 134 154 157 179 192 194-195 227 257 266 274 278 326 355- 357 437 446 498 503 516 532 846

\*The 16 tissue/mRNAs and their vendor sources are as follows: 1) Normal adult brain mRNA (Invitrogen), 2) Normal adult kidney mRNA (Invitrogen), 3) Normal fetal brain mRNA (Invitrogen), 4) Normal adult liver mRNA (Invitrogen), 5) Normal fetal kidney mRNA (Invitrogen), 6) Normal fetal liver mRNA (Invitrogen), 7) normal fetal skin mRNA (Invitrogen), 8) human adrenal gland mRNA (Clontech), 9) Human bone marrow mRNA (Clontech), 10) Human leukemia lymphoblastic mRNA (Clontech), 11) Human thymus mRNA (Clontech), 12) human lymph node mRNA (Clontech), 13) human so\spinal cord mRNA (Clontech), 14)

Table 1  
125

human thyroid mRNA (Clontech), 15) human	esophagus mRNA (BioChain), 16) human
conceptional umbilical cord mRNA (BioChain).	

Table 2A  
126

SEQ ID	Hit ID	Species	Description	S score	% Identity
912	gi12803215	Homo sapiens	glutathione S-transferase theta 2, clone MGC:744 IMAGE:3164017, mRNA, complete cds.	1263	100
912	gi601918	Homo sapiens	glutathione S-transferase theta 2 (GSTT2) mRNA, complete cds.	1263	100
912	gi9937244	Homo sapiens	glutathione S-transferase theta 2 (GSTT2) and glutathione S-transferase theta 1 (GSTT1) genes, complete cds.	1259	99
913	gi13872813	Homo sapiens	partial mRNA for fibulin-6 (FIBL-6 gene).	4548	93
913	gi14575679	Homo sapiens	hemiscentin mRNA, complete cds.	4542	93
913	gi14041957	Homo sapiens	cDNA FLJ14438 fis, clone HEMBB1000317, weakly similar to FIBULIN-1, ISOFORM D PRECURSOR.	2795	85
914	gi13543645	Homo sapiens	D-dopachrome tautomerase, clone MGC:14637 IMAGE:4082777, mRNA, complete cds.	450	94
914	gi15930140	Homo sapiens	Similar to D-dopachrome tautomerase, clone MGC:9158 IMAGE:3902943, mRNA, complete cds.	450	94
914	gi2352915	Homo sapiens	D-dopachrome tautomerase (DDT) gene, exon 3 and complete cds.	450	94
915	gi13543645	Homo sapiens	D-dopachrome tautomerase, clone MGC:14637 IMAGE:4082777, mRNA, complete cds.	537	86
915	gi15930140	Homo sapiens	Similar to D-dopachrome tautomerase, clone MGC:9158 IMAGE:3902943, mRNA, complete cds.	537	86
915	gi2352915	Homo sapiens	D-dopachrome tautomerase (DDT) gene, exon 3 and complete cds.	537	86
916	gi4190954	Homo sapiens	gene for hepatocyte growth factor activator, complete cds.	2089	100
916	gi219681	Homo sapiens	Human mRNA for hepatocyte growth factor (HGF) activator precursor, complete cds.	2089	100
916	AAR89197	Homo sapiens	Human hepatocellular growth factor single chain precursor protein.	2089	100
917	gi5441937	Homo sapiens	BAC clone CTB-15P3 from 7q22-q31.2, complete sequence.	8964	100
917	AAY15457	Homo sapiens	Human laminin beta 4 protein.	6164	96
917	AAY15459	Homo sapiens	SEQ ID 5 of WO9919347.	6046	97
918	gi10998440	Mus musculus	EGF-related protein SCUBE1	2116	52
918	gi8052320	Mus musculus	Cegp1 protein	1124	79
918	AAY07735	Homo sapiens	Human breast-specific BS200 protein.	1125	79
919	gi13543621	Homo sapiens	parathyroid hormone-like hormone, clone MGC:14611 IMAGE:4050706, mRNA, complete cds.	905	100
919	gi190712	Homo sapiens	Human parathyroid hormone-like protein (PLP) gene, exon 4, clones lambda-PLPg(1,3,7-2).	905	100

Table 2A

127

SEQ ID	Hit ID	Species	Description	S score	% Identity
919	gi190713	Homo sapiens	Human parathyroid hormone-like protein (PLP) gene, exon 5, clones lambda-PLPg(1,3,7-2).	905	100
920	gi3449294	Rattus norvegicus	MEGF6	2229	78
920	AAAY72715	Homo sapiens	HFICU08 clone human attractin-like protein.	554	43
920	AAG75479	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6243.	518	43
921	gi1296610	Homo sapiens	H.sapiens mRNA for chemokine CC-2 and CC-1.	346	100
921	gi1004267	Homo sapiens	H.sapiens gene for chemokine HCC-1.	346	100
921	gi1004269	Homo sapiens	H.sapiens mRNA for chemokine HCC-1.	346	100
922	gi35330	Homo sapiens	H.sapiens mRNA for procarboxypeptidase A1.	1126	61
922	AAR97618	Homo sapiens	Human carboxypeptidase A1.	1126	61
922	AAAY28915	Homo sapiens	Human regulatory protein HRGP-1.	1126	61
923	gi790817	Homo sapiens	Human microfibril-associated glycoprotein 4 (MFAP4) mRNA, 3' end of cds.	1198	99
923	AAAY42563	Homo sapiens	Human microfibril-associated glycoprotein 4 splice variant (MAG4V).	1197	100
923	AAAY85177	Homo sapiens	Microfibril associated glycoprotein 4 splice variant protein sequence.	1197	100
924	AAAY06940	Homo sapiens	Human secretory protein ZSIG-11.	1648	100
924	AAAY66650	Homo sapiens	Membrane-bound protein PRO536.	1648	100
924	AAAY50944	Homo sapiens	Human adult heart cDNA clone vfl_1 derived protein.	1648	100
925	gi13528981	Homo sapiens	apolipoprotein A-II, clone MGC:12334 IMAGE:3934476, mRNA, complete cds.	438	90
925	gi28748	Homo sapiens	Human mRNA for apolipoprotein AII precursor.	438	90
925	gi296633	Homo sapiens	Human DNA for apolipoprotein A-II.	438	90
926	AAAY76156	Homo sapiens	Human secreted protein encoded by gene 33.	419	90
926	AAE04850	Homo sapiens	Human SGP014 phosphatase polypeptide related exon 4.	63	33
926	gi599904	Albinaria turrita	ATPase subunit 8	40	39
927	gi13097252	Homo sapiens	Similar to FK506 binding protein 2 (13 kDa), clone MGC:5177 IMAGE:3445148, mRNA, complete cds.	640	91
927	gi337370	Homo sapiens	Human rapamycin- and FK506-binding protein, complete cds.	640	91
927	AAQ31004_aa1	Homo sapiens	hRFBKP cDNA.	636	90
928	AAB53360	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:900.	230	95
928	AAG73789	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:4553.	230	95

Table 2A  
128

SEQ ID	Hit ID	Species	Description	S score	% Identity
928	AAAY86330	Homo sapiens	Human secreted protein HKGCO27, SEQ ID NO:245.	60	66
929	gi191613	Mus musculus	acetylcholine receptor	55	43
929	AAAY07766	Homo sapiens	Human secreted protein fragment encoded from gene 23.	55	40
930	gi15157181	Agrobacterium tumefaciens	AGR_C_3718p	1148	50
930	gi14022240	Mesorhizobium loti	probable D-lactate dehydrogenase	1106	47
930	gi15075160	Sinorhizobium meliloti	PUTATIVE D-LACTATE DEHYDROGENASE (CYTOCHROME) PROTEIN	1097	48
931	AAW64548	Homo sapiens	Human epidermoid carcinoma cell line KB clone HP10179 protein.	231	100
931	AAW78146	Homo sapiens	Human secreted protein encoded by gene 21 clone HWTAZ75.	231	100
931	gi7106778	Homo sapiens	HSPC194	151	75
932	AAAY24793	Homo sapiens	Human secreted protein yc2 1.	1107	99
932	gi3858883	Acanthamoeba castellanii	myosin I heavy chain kinase	143	30
932	gi13751817	Leishmania major	C2 domain protein	126	32
933	gi15028816	Homo sapiens	mRNA for beta-1,3-galactosyltransferase b3Gal-T8.	172	28
933	gi8927164	Mus musculus	core1 UDP-galactose:N-acetylgalactosamine-alpha-R beta 1,3-galactosyltransferase	173	28
933	gi8927166	Rattus norvegicus	core1 UDP-galactose:N-acetylgalactosamine-alpha-R beta 1,3-galactosyltransferase	173	26
934	AAAY73383	Homo sapiens	HTRM clone 2280456 protein sequence.	1571	100
934	gi15929192	Homo sapiens	clone MGC:9522 IMAGE:3909690, mRNA, complete cds.	1487	99
934	AAG75282	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6046.	619	99
935	gi15705411	Homo sapiens	peptidoglycan recognition protein L precursor (PGLYRP) mRNA, complete cds.	3041	99
935	AAAY72664	Homo sapiens	Murine peptidoglycan recognition protein-related liver protein.	1484	76
935	gi6651239	Mus musculus	TAGL-alpha	1471	76
936	AAB24450	Homo sapiens	Human secreted protein sequence encoded by gene 14 SEQ ID NO:75.	361	97
936	AAB49409	Homo sapiens	hCRF1-TM8 construct peptide.	39	69
936	gi12619689	Conus ventricosus	conotoxin scaffold VI/VII precursor	58	21
937	gi12804373	Homo sapiens	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide, clone MGC:675 IMAGE:3543571, mRNA, complete cds.	1094	94
937	gi1711232	Homo sapiens	Human DNA for 14-3-3 protein eta chain, exon2 and complete cds.	1094	94



Table 2A

129

SEQ ID	Hit ID	Species	Description	S score	% Identity
937	gi531590	Homo sapiens	H. sapiens gene for 14-3-3 protein.	1094	94
938	gi12804681	Homo sapiens	S100 calcium-binding protein, beta (neural), clone MGC:1323 IMAGE:3543825, mRNA, complete cds.	479	100
938	gi337730	Homo sapiens	Human S100 protein beta-subunit gene, exon 3.	479	100
938	gi16118441	Oryctolagus cuniculus	S-100 calcium-binding protein beta subunit	479	100
939	AAW75082	Homo sapiens	Human secreted protein encoded by gene 26 clone HTLEV12.	392	92
939	AAG00461	Homo sapiens	Human secreted protein, SEQ ID NO: 4542.	72	50
939	AAV86322	Homo sapiens	Human secreted protein HTLFG05, SEQ ID NO:237.	72	50
941	gi9280025	Macaca fascicularis	Nogo receptor	804	49
941	gi15080005	Homo sapiens	nogo receptor, clone MGC:19831 IMAGE:4040540, mRNA, complete cds.	792	49
941	gi12407653	Homo sapiens	Nogo receptor mRNA, complete cds.	792	49
942	AAB25674	Homo sapiens	Human secreted protein sequence encoded by gene 10 SEQ ID NO:63.	768	98
942	gi14603247	Homo sapiens	Similar to RIKEN cDNA 5730409G15 gene, clone MGC:19636 IMAGE:2822323, mRNA, complete cds.	238	90
942	AAB36613	Homo sapiens	Human FLEXHT-35 protein sequence SEQ ID NO:35.	238	90
943	AAV85678	Homo sapiens	Human kidney disease associated protein SEQ ID 10.	751	98
943	gi3127193	Rattus norvegicus	kidney-specific protein	686	75
943	gi5019275	Bos taurus	xenobiotic/medium-chain fatty acid:CoA ligase form XL-III	474	54
944	gi1197499	Homo sapiens	H.sapiens gene for C1 inhibitor exon 2 (and joined CDS).	2527	100
944	gi29535	Homo sapiens	Human gene for C1-inhibitor.	2527	100
944	gi15029894	Homo sapiens	serine (or cysteine) proteinase inhibitor, clade G (C1 inhibitor), member 1, clone MGC:17091 IMAGE:4150091, mRNA, complete cds.	2524	99
945	gi15157854	Agrobacterium tumefaciens	AGR_C_4799p	218	38
945	gi15076099	Sinorhizobium meliloti	HYPOTHETICAL PROTEIN	233	36
945	gi9951121	Pseudomonas aeruginosa	ribosomal protein L11 methyltransferase	116	38
946	AAB03948	Homo sapiens	Human mesenchymal stem cell polypeptide.	462	98
946	AAB64909	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:87.	166	63
946	gi1531983	Homo sapiens	H.sapiens mRNA for CC-	77	31

Table 2A

130

SEQ ID	Hit ID	Species	Description	S score	% Identity
			chemokine, eotaxin variant (clone 53).		
947	gi13958036	Homo sapiens	FYVE-finger protein EIP1 mRNA, complete cds.	2917	97
947	AAAY29861	Homo sapiens	Human secreted protein clone cb98_4.	2910	96
947	gi11558044	Mus musculus	FYVE-finger containing protein	2752	91
948	AAW75000	Homo sapiens	Human secreted protein encoded by gene 146 clone HSNK17.	49	29
948	gi10176829	Arabidopsis thaliana	gene_id:MBB18.16~	79	32
949	gi13279266	Homo sapiens	clone MGC:10946 IMAGE:3631700, mRNA, complete cds.	466	100
949	AAG81435	Homo sapiens	Human AFP protein sequence SEQ ID NO:388.	466	100
949	AAB03208	Homo sapiens	Human gene 7 encoded secreted protein HNTDL21, SEQ ID NO:58.	466	100
950	AAAY16787	Homo sapiens	Human secreted protein (clone dy41_2).	557	100
950	gi6682818	Sus scrofa	lectin-like oxidized LDL receptor-1	235	29
950	AAW52837	Homo sapiens	Human C-type lectin MCTL.	206	34
951	AAAY41720	Homo sapiens	Human PRO792 protein sequence.	1140	87
951	AAB44276	Homo sapiens	Human PRO792 (UNQ431) protein sequence SEQ ID NO:231.	1140	87
951	AAB24055	Homo sapiens	Human PRO792 protein sequence SEQ ID NO:31.	1140	87
952	gi7670746	Homo sapiens	UDP-glucose:glycoprotein glucosyltransferase 1 precursor, mRNA, complete cds.	7968	98
952	gi13275621	synthetic construct	Rat RUGT	7313	91
952	gi7677176	Rattus norvegicus	UDP-glucose glycoprotein:glucosyltransferase precursor	7313	91
953	AAB64390	Homo sapiens	Amino acid sequence of human intracellular signalling molecule INTRA22.	409	51
953	AAG00975	Homo sapiens	Human secreted protein, SEQ ID NO: 5056.	212	46
953	gi2623224	Mus musculus	FK-506 binding protein homolog	118	30
954	gi1778061	Saccharomyces cerevisiae	transcription/repair factor TFIIH subunit Tfb3	83	31
954	gi927727	Saccharomyces cerevisiae	Tfb3p: TFIIH subunit Tfb3; YDR460W	83	31
954	gi3372804	Rattus norvegicus	focal adhesion kinase-related protein	65	25
955	AAG81399	Homo sapiens	Human AFP protein sequence SEQ ID NO:316.	1910	100
955	AAB61421	Homo sapiens	Human TANGO 300 protein.	1904	99
955	AAB23618	Homo sapiens	Human secreted protein SEQ ID NO: 36.	1902	99
956	gi296532	Homo sapiens	H.sapiens mRNA for I beta 1-6 N-acetylglucosaminyltransferase.	1043	66

Table 2A

131

SEQ ID	Hit ID	Species	Description	S score	% Identity
956	gi1315909	Homo sapiens	Human beta-1,6-N-acetylglucosaminyltransferase (IGnT) gene, exon 3, complete cds.	1043	66
956	gi307298	Homo sapiens	Human I beta 1-6 N-acetylglucosaminyltransferase mRNA, complete cds.	1043	66
957	gi2065165	Homo sapiens	H.sapiens mRNA for extracellular matrix protein collagen type XIV, N-terminus.	776	93
957	AAG00322	Homo sapiens	Human secreted protein, SEQ ID NO: 4403.	482	98
957	gi288875	Gallus gallus	collagen XIV	467	68
958	gi15530201	Homo sapiens	Similar to matrix metalloproteinase 1 (interstitial collagenase), clone MGC:10479 IMAGE:3834572, mRNA, complete cds.	2381	93
958	gi30126	Homo sapiens	H.sapiens mRNA for type I interstitial collagenase.	2381	93
958	gi1688258	Homo sapiens	Human collagenase and stromelysin genes, complete cds, and metalloelastase gene, partial cds.	2381	93
959	AAV25868	Homo sapiens	Human secreted protein fragment encoded from gene 57.	993	99
959	gi4008584	Rattus norvegicus	huntingtin	61	66
959	gi6006615	Saimiri sciureus	Alanine Glyoxylate Aminotransferase	71	31
960	AAB58288	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 626.	474	85
960	gi13378304	Homo sapiens	dynein-associated protein HKM23 (km23) mRNA, complete cds.	473	98
960	gi12803327	Homo sapiens	HSPC162 protein, clone MGC:773 IMAGE:3347555, mRNA, complete cds.	473	98
961	gi930078	Homo sapiens	Human Kox16 mRNA for zinc finger protein, partial.	56	34
961	AAE01364	Homo sapiens	Human gene 13 encoded secreted protein HDPIW06, SEQ ID NO:86.	54	43
961	AAE01398	Homo sapiens	Human gene 13 encoded secreted protein HDPIW06, SEQ ID NO:120.	54	43
962	gi9956936	Mus musculus	Su(var)3-9 homolog Suv39h2	1313	84
962	gi10440094	Homo sapiens	cDNA: FLJ23414 fis, clone HEP20704.	1236	100
962	gi13623277	Homo sapiens	suppressor of variegation 3-9 (Drosophila) homolog 1, clone MGC:10376 IMAGE:3945632, mRNA, complete cds.	821	55
963	gi12654023	Homo sapiens	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5 (13kD, B13), clone MGC:5042 IMAGE:3453236, mRNA, complete cds.	432	90
963	gi1373173	Homo sapiens	NADH:ubiquinone oxidoreductase	432	90

Table 2A

132

SEQ ID	Hit ID	Species	Description	S score	% Identity
			subunit B13 (B13) mRNA, complete cds.		
963	gi1698653	Homo sapiens	Human NADH:ubiquinone oxidoreductase subunit B13 mRNA, complete cds.	432	90
964	gi37347	Homo sapiens	Human mRNA for T-cell rearranging gamma gene (TRG) V(g)8-J(g)2-C(g)2.	1645	94
964	gi339407	Homo sapiens	Human T-cell receptor Ti rearranged gamma-chain mRNA V-J-C region, complete cds.	1617	93
964	gi37018	Homo sapiens	Human mRNA for T-cell receptor gamma-chain.	1439	87
965	AAE06606	Homo sapiens	Human protein having hydrophobic domain, HP10794.	566	96
965	gi6996628	Triticum aestivum	phenylalanine ammonia lyase	72	46
965	gi12667442	Rattus norvegicus	synaptotagmin VIIT1	75	34
966	gi13654639	Bos taurus	D-glucuronyl C5 epimerase	3165	97
966	gi13442978	Mus musculus	D-glucuronyl C5-epimerase	3142	95
966	gi11935177	Mus musculus	heparin/heparan sulfate:glucuronic acid C5 epimerase	3137	95
967	AAG71875	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1556.	709	86
967	AAG71443	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1124.	692	79
967	AAG71816	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1497.	684	82
968	gi12655133	Homo sapiens	Similar to CGI-63 protein, clone MGC:708 IMAGE:3139846, mRNA, complete cds.	1918	100
968	AAB73679	Homo sapiens	Human oxidoreductase protein ORP-12.	1918	100
968	gi4929595	Homo sapiens	CGI-63 protein mRNA, complete cds.	1874	98
969	gi6062874	Homo sapiens	candidate tumor suppressor protein DICE1 mRNA, complete cds.	1313	81
969	AAAY15344	Homo sapiens	Tumour suppressor protein del-27.	1313	81
969	AAAY28995	Homo sapiens	Tumour suppressor Del-27 protein sequence.	1313	81
970	gi387011	Homo sapiens	Human pyruvate dehydrogenase E1-alpha subunit mRNA, cds.	2187	99
970	gi12803199	Homo sapiens	pyruvate dehydrogenase (lipoamide) alpha 1, clone MGC:8609 IMAGE:2961286, mRNA, complete cds.	2049	100
970	gi35379	Homo sapiens	Human mRNA for brain pyruvate dehydrogenase (EC 1.2.4.1).	2049	100
971	gi2275569	Homo sapiens	T cell receptor beta locus, TCRBV6S4A1 to TCRBV8S1 region.	660	99
971	gi2104755	Homo sapiens	T cell receptor V-beta 23 (TCRBV) gene, partial cds.	660	99
971	gi495404	Pan troglodytes	T cell receptor beta chain	657	88
972	AAG89141	Homo sapiens	Human secreted protein, SEQ ID	1272	100

Table 2A  
133

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NO: 261.		
972	gi12053273	Homo sapiens	mRNA; cDNA DKFZp434K1172 (from clone DKFZp434K1172); complete cds.	1271	99
972	gi292836	Homo sapiens	Human trichohyalin (TRHY) gene, complete cds.	146	29
973	gi12654633	Homo sapiens	protein inhibitor of activated STAT3, clone MGC:1417 IMAGE:3528679, mRNA, complete cds.	2708	96
973	gi4996563	Homo sapiens	PIAS3 mRNA for protein inhibitor of activated STAT3, complete cds.	2708	96
973	AAE02937	Homo sapiens	Human TFRP protein.	2708	96
974	gi15099957	Homo sapiens	diacylglycerol acyltransferase 2-like protein mRNA, complete cds.	856	79
974	gi15099955	Mus musculus	diacylglycerol acyltransferase 2-like protein	676	59
974	AAV94889	Homo sapiens	Human protein clone HP02485.	578	56
975	gi7022243	Homo sapiens	cDNA FLJ10300 fis, clone NT2RM2000030.	1432	100
975	AAB92669	Homo sapiens	Human protein sequence SEQ ID NO:11033.	1432	100
975	gi3834427	Drosophila melanogaster	cytoplasmic dynein intermediate chain isoform DIC1c	134	24
976	gi12803965	Homo sapiens	clone MGC:4294 IMAGE:3636069, mRNA, complete cds.	519	100
976	gi15919713	Homo sapiens	isolate NSB11-3-K3-A10 immunoglobulin kappa chain variable region gene, partial cds.	62	32
976	gi15919629	Homo sapiens	isolate NSB11-4-K3-B6 immunoglobulin kappa chain variable region gene, partial cds.	60	33
977	gi1209685	Homo sapiens	Human salivary peroxidase mRNA, complete cds.	3408	99
977	gi163307	Bos taurus	lactoperoxidase	2906	77
977	gi11990122	Camelus dromedarius	peroxidase	2882	84
978	AAV01603	Homo sapiens	Amino acid sequence of the human defensin (Def-X) protein.	501	97
978	gi29735	Homo sapiens	H.sapiens mRNA for corticostatin HP-4 precursor.	214	46
978	gi665927	Homo sapiens	Human corticostatin/defensin HP-4 precursor gene, complete cds.	214	46
979	AAG81415	Homo sapiens	Human AFP protein sequence SEQ ID NO:348.	848	100
979	gi11559416	Oryctolagus cuniculus	NADPH-dependent retinol dehydrogenase/reductase	75	31
980	AAR94422	Homo sapiens	Bactericidal/permeability increasing peptide (BPI.245).	63	56
980	AAR76424	Homo sapiens	Bacterial permeability-increasing peptide BPI.245.	63	56
980	AAW63485	Homo sapiens	Human BPI protein derived peptide XMP.245.	63	56
981	gi4309953	Homo sapiens	BAC clone RP11-121A8 from 7p14-p13, complete sequence.	951	95

Table 2A

134

SEQ ID	Hit ID	Species	Description	S score	% Identity
981	gi296680	Homo sapiens	Human germline TCR-gamma gene variable region V-gamma 5.	570	85
981	gi4309952	Homo sapiens	BAC clone RP11-121A8 from 7p14-p13, complete sequence.	544	68
982	gi1223888	synthetic construct	T cell receptor alpha chain	1070	79
982	gi338766	Homo sapiens	Human T-cell receptor rearranged alpha-chain V-region (V-D-J) mRNA, complete cds.	994	73
982	gi3089419	Homo sapiens	SSC11 rearranged T cell receptor alpha chain (TCRAV17) gene, complete cds.	976	71
983	gi14249942	Homo sapiens	Similar to RIKEN cDNA 061008P16 gene, clone MGC:15937 IMAGE:3537224, mRNA, complete cds.	309	46
983	AAB73512	Homo sapiens	Human transferase HTFS-19, SEQ ID NO:19.	309	46
983	gi2828262	Bos taurus	aralkyl acyl-CoA:amino acid N-acyltransferase	285	42
984	AAG71251	Homo sapiens	Human gene 9-encoded secreted protein HMSDL37, SEQ ID NO:99.	311	90
984	AAG71286	Homo sapiens	Human gene 9-encoded secreted protein HMSDL37, SEQ ID NO:135.	311	90
984	gi13096922	Mus musculus	Similar to nadrin	76	44
985	gi4519541	Mus musculus	thrombospondin type 1 domain	1299	88
985	gi13625176	Homo sapiens	clone 1 thrombospondin mRNA, complete cds.	646	46
985	AAW85607	Homo sapiens	Secreted protein clone da228_6.	646	46
986	AAG81417	Homo sapiens	Human AFP protein sequence SEQ ID NO:352.	287	100
986	gi176558	Aotus trivirgatus	involucrin (large allele)	53	40
986	AAG77243	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:8009.	36	70
987	gi2104856	Rattus norvegicus	platelet glycoprotein V	551	36
987	gi6449037	Mus musculus	platelet glycoprotein V	550	36
987	gi312502	Homo sapiens	H.sapiens GPV gene encoding platelet glycoprotein V precursor.	486	34
988	gi12803851	Homo sapiens	nucleotide binding protein 2 (E.coli MinD like), clone MGC:3473 IMAGE:3633393, mRNA, complete cds.	1385	100
988	gi14124958	Homo sapiens	nucleotide binding protein 2 (E.coli MinD like), clone MGC:15834 IMAGE:3507768, mRNA, complete cds.	1385	100
988	AAB59022	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 730.	1385	100
989	gi9864185	Drosophila melanogaster	Crossveinless 2	1031	35
989	gi7768636	Xenopus laevis	Kielin	911	31
989	gi9887910	Oryctolagus cuniculus	zonadhesin precursor	540	31

Table 2A  
135

SEQ ID	Hit ID	Species	Description	S score	% Identity
990	gi13161011	Homo sapiens	sclerostin gene, complete cds.	417	92
990	gi13161020	Homo sapiens	sclerostin mRNA, complete cds.	417	92
990	gi13236418	Homo sapiens	SOST (SOST) mRNA, complete cds.	417	92
991	gi335148	Tacaribe virus	P-11 protein	69	39
991	gi5596437	Petunia x hybrida	Stig1	58	30
992	gi4490538	Homo sapiens	Human U266 rearranged DNA for lambda-immunoglobulin light chain.	933	81
992	gi33746	Homo sapiens	Human rearranged immunoglobulin lambda light chain mRNA.	920	80
992	AAB36212	Homo sapiens	Human immune system associated protein HISAP-10.	913	81
993	AAG03466	Homo sapiens	Human secreted protein, SEQ ID NO: 7547.	104	31
993	gi259589	Cercopithecine herpesvirus 1	glycoprotein J; gJ	88	42
993	gi3514059	Cercopithecine herpesvirus 1	glycoprotein gJ	87	42
994	gi203246	Rattus norvegicus	cell adhesion-like molecule	1739	97
994	gi514374	Homo sapiens	Human (clone pHOM) opioid-binding cell adhesion molecule mRNA, complete cds.	1666	94
994	gi586	Bos taurus	put. pre-OPCAM (AA 1 - 345)	1643	93
995	AAB88408	Homo sapiens	Human membrane or secretory protein clone PSEC0164.	1358	92
995	AAY13392	Homo sapiens	Amino acid sequence of protein PRO328.	1355	100
995	AAB01373	Homo sapiens	Neuron-associated protein.	1355	100
996	AAB85144	Homo sapiens	Human NKCR polypeptide (clone ID HMSOM53).	982	89
996	gi31332	Homo sapiens	Human mRNA for high affinity Fc receptor (FcRI).	364	42
996	AAY96226	Homo sapiens	Human high affinity Fc receptor, FcgammaRI.	364	42
997	gi14165486	Homo sapiens	clone MGC:15041 IMAGE:3831657, mRNA, complete cds.	70	37
997	gi9909361	Homo sapiens	sphingosine kinase (SPHK) mRNA, complete cds.	70	37
997	gi8132868	Homo sapiens	sphingosine kinase-1 mRNA, complete cds.	70	37
998	AAB56802	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1380.	417	100
998	gi2983321	Aquifex aeolicus	cytochrome c-type biogenesis protein	61	36
998	gi6941880	Escherichia coli	flagellin	63	35
999	AAB08732	Homo sapiens	Amino acid sequence of a human OLD-35 polypeptide.	1375	87
999	gi14042110	Homo sapiens	cDNA FLJ14531 fis, clone NT2RM2000371, weakly similar to POLYRIBONUCLEOTIDE	1265	88

Table 2A  
136

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NUCLEOTIDYLTRANSFERASE (EC 2.7.7.8).		
999	AAB92684	Homo sapiens	Human protein sequence SEQ ID NO:11065.	1265	88
1000	gi199582	Mus musculus	B(2)-microglobulin	616	95
1000	gi50105	Mus musculus	beta2-microglobulin precursor (aa -20 to 99)	614	95
1000	gi199576	Mus musculus	B(2)-microglobulin	609	94
1001	gi7023273	Homo sapiens	cDNA FLJ10922 fis, clone OVARC1000420.	672	99
1001	AAB93335	Homo sapiens	Human protein sequence SEQ ID NO:12441.	672	99
1001	AAM06337	Homo sapiens	Human foetal protein, SEQ ID NO: 68.	328	75
1002	gi15559608	Homo sapiens	Similar to zinc finger protein 16 (KOX 9), clone MGC:20886 IMAGE:4549240, mRNA, complete cds.	2079	100
1002	gi55471	Mus musculus	Zfp-29	725	59
1002	gi6409345	Homo sapiens	zinc finger protein ZNF180 (ZNF180) mRNA, complete cds.	730	54
1003	gi488555	Homo sapiens	Human zinc finger protein ZNF135 mRNA, complete cds.	1664	60
1003	gi1769491	Homo sapiens	Human kruppel-related zinc finger protein (ZNF184) mRNA, partial cds.	1633	49
1003	gi10436789	Homo sapiens	cDNA FLJ14345 fis, clone THYRO1001189, weakly similar to ZINC FINGER PROTEIN 91.	1619	57
1005	AAB23641	Homo sapiens	Human secreted protein SEQ ID NO: 97.	609	100
1005	AAE01594	Homo sapiens	Human gene 12 encoded secreted protein HHSGW69, SEQ ID NO:144.	93	35
1005	gi5668598	Homo sapiens	Wiskott-Aldrich syndrome protein interacting protein (WASPIP) mRNA, partial cds.	92	35
1006	gi11493473	Homo sapiens	PRO2225	163	67
1006	gi9654986	Vibrio cholerae	autoinducer-2 production protein	64	33
1006	AAG02328	Homo sapiens	Human secreted protein, SEQ ID NO: 6409.	58	45
1008	AAM00955	Homo sapiens	Human bone marrow protein, SEQ ID NO: 431.	704	100
1008	gi4902661	Homo sapiens	Novel human gene mapping to chromosome 22.	635	100
1008	gi12052896	Homo sapiens	mRNA; cDNA DKFZp564F1978 (from clone DKFZp564F1978); complete cds.	635	100
1009	AAE01420	Homo sapiens	Human secreted protein fragment, SEQ ID NO:144.	612	100
1009	gi4028545	Homo sapiens	LIM domain only 7 (LMO7) gene, exon T and alternative splice products, partial cds.	65	44
1009	gi1620061	Paramecium bursaria Chlorella virus 1	a389R	66	46



Table 2A  
137

SEQ ID	Hit ID	Species	Description	S score	% Identity
1010	gi9408382	Raja eglanteria	eos	79	26
1010	gi2062744	Xenopus laevis/gilli	Ikaros homolog	63	27
1010	gi1199604	Homo sapiens	Human zinc finger protein C2H2-25 mRNA, complete cds.	97	27
1011	gi13097207	Homo sapiens	ribosomal protein, large, P1, clone MGC:5215 IMAGE:2900846, mRNA, complete cds.	332	100
1011	gi14043204	Homo sapiens	ribosomal protein, large, P1, clone MGC:15616 IMAGE:3343021, mRNA, complete cds.	332	100
1011	gi190234	Homo sapiens	Human acidic ribosomal phosphoprotein P1 mRNA, complete cds.	332	100
1012	AAW88457	Homo sapiens	Human lysophospholipase IHLF.	1113	93
1012	AAV64648	Homo sapiens	Human lysophospholipase homology protein.	825	100
1012	gi10303289	Neurospora crassa	related to lysophospholipase	201	34
1013	gi14043417	Homo sapiens	clone IMAGE:3953868, mRNA, partial cds.	919	98
1013	gi15080096	Homo sapiens	clone MGC:20451 IMAGE:3830864, mRNA, complete cds.	919	98
1013	AAG81374	Homo sapiens	Human AFP protein sequence SEQ ID NO:266.	919	98
1014	gi13543427	Homo sapiens	Similar to RIKEN cDNA 5730469M10 gene, clone MGC:4248 IMAGE:3010078, mRNA, complete cds.	191	36
1014	AAW67858	Homo sapiens	Human secreted protein encoded by gene 52 clone HAUCC47.	191	36
1014	AAV66746	Homo sapiens	Membrane-bound protein PRO1198.	191	36
1015	AAB56791	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1369.	496	98
1015	gi2984030	Aquifex aeolicus	phosphofructokinase	70	25
1015	gi1388150	Haemophilus ducreyi	ribosomal protein L31	51	45
1016	gi29509	Homo sapiens	Human BTG1 mRNA.	655	100
1016	gi50188	Mus musculus	btg1	655	100
1016	gi293306	Mus musculus	B-cell translocation gene-1 protein	655	100
1017	gi13623633	Homo sapiens	clone MGC:13105 IMAGE:3957973, mRNA, complete cds.	2984	100
1017	AAB81188	Homo sapiens	Human zinc finger protein 52 (ZFP-52).	2389	98
1017	AAB95368	Homo sapiens	Human protein sequence SEQ ID NO:17684.	1712	100
1018	gi12002127	Homo sapiens	CCK1 protein (CCK1) mRNA, complete cds.	55	45
1018	gi9392591	Homo sapiens	CC chemokine CCL28 (SCYA28) mRNA, complete cds.	55	45
1018	gi10312152	Homo sapiens	mucosae-associated epithelial chemokine mRNA, complete cds.	55	45

Table 2A  
138

SEQ ID	Hit ID	Species	Description	S score	% Identity
1019	AAB90816	Homo sapiens	Human shear stress-response protein SEQ ID NO: 140.	365	100
1019	gi10039341	Dugesia dorocephala	Vasa-related protein PIVAS1	67	45
1019	gi15158371	Agrobacterium tumefaciens	AGR_L_42Gmp	64	30
1020	gi1377897	Homo sapiens	heart protein (FHL-2) mRNA, complete cds.	1631	100
1020	gi1160932	Homo sapiens	(clone 35.3) DRAL mRNA, complete cds.	1625	99
1020	gi7209525	Homo sapiens	gene for DRAL/Slim3/FHL2, exon 5 and complete cds.	1625	99
1021	gi7770259	Homo sapiens	PRO2975	794	100
1021	gi1142588	Trypanosoma brucei	CR3	55	40
1021	gi1237130	Escherichia coli	O antigen polymerase	89	28
1022	AAG81348	Homo sapiens	Human AFP protein sequence SEQ ID NO:214.	530	63
1022	gi4206763	Arabidopsis thaliana	cell wall-plasma membrane linker protein homolog	119	30
1022	gi2623666	Canis familiaris	dentatorubro-pallidolusian atrophy protein	101	30
1023	AAG01390	Homo sapiens	Human secreted protein, SEQ ID NO: 5471.	297	100
1023	AAB57280	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1858.	66	34
1023	gi456681	Pseudorabies virus	helicase	80	26
1024	gi1017722	Homo sapiens	Human repressor transcriptional factor (ZNF85) mRNA, complete cds.	1344	58
1024	gi14348591	Homo sapiens	KRAB zinc finger protein (KR19) mRNA, complete cds.	1342	58
1024	gi4454678	Homo sapiens	zinc finger protein 4	1321	59
1025	gi165680	Oryctolagus cuniculus	recombination activating protein	2666	91
1025	gi2576246	Mus musculus	RAG-2 protein	2594	88
1025	gi15809414	Mormoops blainvillii	recombination activator protein 2	2387	91
1026	AAG75278	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6042.	1061	95
1026	gi211896	Gallus gallus	h-caldesmon	153	29
1026	gi600255	Gallus gallus	caldesmon	153	29
1027	AAY87341	Homo sapiens	Human signal peptide containing protein HSPP-118 SEQ ID NO:118.	699	99
1027	AAY12929	Homo sapiens	Amino acid sequence of a human secreted peptide.	66	51
1027	gi930084	Homo sapiens	Human Kox21 mRNA for zinc finger protein, partial.	41	31
1028	gi173324	Candida glabrata	metallothionein II	37	71
1028	gi173326	Candida glabrata	metallothionein (MTII)	37	71
1028	gi173328	Candida glabrata	metallothionein II	37	71
1029	AAE02058	Homo sapiens	Human four disulfide core domain (FDCD)-containing protein.	598	44

Table 2A

139

SEQ ID	Hit ID	Species	Description	S score	% Identity
1029	gi12655452	Homo sapiens	mRNA for keratin associated protein 4.7 (KRTAP4.7 gene).	557	46
1029	gi12655456	Homo sapiens	partial mRNA for keratin associated protein 4.9 (KRTAP4.9 gene).	540	42
1030	gi10639287	Thermoplasma acidophilum	amino acid transporter related protein	66	29
1030	gi14325303	Thermoplasma volcanium	amino acid transporter	64	28
1030	gi15341701	Homo sapiens	clone MGC:19805 IMAGE:3939974, mRNA, complete cds.	60	34
1031	gi12653801	Homo sapiens	peptidylprolyl isomerase A (cyclophilin A), clone MGC:2351 IMAGE:3349335, mRNA, complete cds.	820	93
1031	gi12804335	Homo sapiens	clone IMAGE:2823490, mRNA, partial cds.	820	93
1031	gi13529080	Homo sapiens	peptidylprolyl isomerase A (cyclophilin A), clone MGC:12404 IMAGE:3935025, mRNA, complete cds.	820	93
1032	gi207621	Rattus norvegicus	uromodulin	98	36
1032	gi912817	Rattus sp.	Tamm-Horsfall protein; THP	98	36
1032	gi602768	Mytilus galloprovincialis	adhesive plaque matrix protein precursor	86	29
1033	AAG03055	Homo sapiens	Human secreted protein, SEQ ID NO: 7136.	269	100
1033	AAY73471	Homo sapiens	Human secreted protein clone yd153_1 protein sequence SEQ ID NO:164.	62	33
1034	AAB38043	Homo sapiens	Fragment of human secreted protein encoded by gene 10 clone HWHGP71.	124	38
1034	gi5305335	Mycobacterium tuberculosis	proline-rich mucin homolog	114	36
1034	gi22599	Arabidopsis thaliana	APG	112	37
1035	gi6467206	Homo sapiens	GIOT-4 mRNA for gonadotropin inducible transcription repressor-4, complete cds.	2004	60
1035	AAY58627	Homo sapiens	Protein regulating gene expression PRGE-20.	1732	56
1035	gi3953593	Mus musculus	Zinc finger protein s11-6	1720	54
1036	AAB95007	Homo sapiens	Human protein sequence SEQ ID NO:16685.	518	86
1036	gi45906	Proteus vulgaris	hlyC protein (AA 1-54)	72	45
1036	AAB56607	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1185.	74	56
1037	AAY27616	Homo sapiens	Human secreted protein encoded by gene No. 50.	562	99
1037	gi12957417	Casuarium bennetti	ATPase subunit 8	62	35
1037	gi332009	Murine leukemia virus	p15-gag protein	57	27

Table 2A

140

SEQ ID	Hit ID	Species	Description	S score	% Identity
1038	gi15822827	Homo sapiens	mRNA for pendrin-like protein 1, complete cds.	1154	39
1038	gi13344999	Homo sapiens	solute carrier family 26 member 6 (SLC26A6) mRNA, complete cds.	1300	37
1038	AAY71067	Homo sapiens	Human membrane transport protein, MTRP-12.	1297	37
1039	AAY57945	Homo sapiens	Human transmembrane protein HTMPN-69.	780	100
1039	AAY76141	Homo sapiens	Human secreted protein encoded by gene 18.	780	100
1039	AAB24037	Homo sapiens	Human PRO1555 protein sequence SEQ ID NO:49.	356	47
1040	AAY59672	Homo sapiens	Secreted protein 108-006-5-0-E6-FL.	553	83
1040	gi10435214	Homo sapiens	cDNA FLJ13263 fis, clone OVARC1000924.	549	82
1040	AAB94543	Homo sapiens	Human protein sequence SEQ ID NO:15290.	549	82
1041	AAY92710	Homo sapiens	Human membrane-associated protein Zsig24.	704	97
1041	AAY87250	Homo sapiens	Human signal peptide containing protein HSPP-27 SEQ ID NO:27.	566	99
1041	AAG00627	Homo sapiens	Human secreted protein, SEQ ID NO: 4708.	260	100
1042	gi14572521	Homo sapiens	NEPH1 (NEPH1) mRNA, complete cds.	1512	51
1042	AAB37996	Homo sapiens	Human secreted protein encoded by gene 13 clone HIBEU15.	1164	92
1042	gi10434261	Homo sapiens	cDNA FLJ12646 fis, clone NT2RM4001987, weakly similar to NEURAL CELL ADHESION MOLECULE 1, LARGE ISOFORM PRECURSOR.	1035	41
1043	gi29806	Homo sapiens	Human mRNA for CD59, an LY-6-like protein regulating complement membrane attack.	710	100
1043	gi825637	Homo sapiens	H.sapiens gene for CD59 protein, exon 2.	710	100
1043	gi29815	Homo sapiens	Human mRNA for CD59 antigen.	710	100
1044	gi6841140	Homo sapiens	HSPC100 mRNA, partial cds.	498	100
1044	gi2828808	Bacillus subtilis	glucose transporter	111	25
1044	gi9106658	Xylella fastidiosa 9a5c	glucose/galactose transporter	140	23
1045	AAB56632	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1210.	3377	99
1045	gi13097708	Homo sapiens	ribophorin II, clone MGC:1817 IMAGE:3546673, mRNA, complete cds.	3152	100
1045	gi5834424	Homo sapiens	RIBIIR gene (partial), exon 1 and joined CDS.	3152	100
1046	gi13182757	Homo sapiens	HTPAP mRNA, complete cds.	598	100
1046	AAG89279	Homo sapiens	Human secreted protein, SEQ ID NO: 399.	598	100
1046	AAB70690	Homo sapiens	Human hDPP protein sequence SEQ ID NO:7.	598	100

Table 2A  
141

SEQ ID	Hit ID	Species	Description	S score	% Identity
1047	gi2276448	Homo sapiens	Human MHC class I HLA-A (HLA-A-0302-new allele) mRNA, complete cds.	1794	93
1047	gi6815812	Homo sapiens	MHC class I antigen heavy chain (HLA-A) mRNA, HLA-A*0302 allele, complete cds.	1794	93
1047	gi1245460	Homo sapiens	Human MHC class I HLA-A allele (HLA-A) mRNA, complete cds.	1786	92
1048	AAB95392	Homo sapiens	Human protein sequence SEQ ID NO:17743.	567	78
1048	AAB29645	Homo sapiens	Human membrane-associated protein HUMAP-2.	548	70
1048	AAB95049	Homo sapiens	Human protein sequence SEQ ID NO:16845.	396	78
1049	gi14017773	Mus musculus	Cg10671-like	1517	96
1049	gi14017764	Mus musculus	CG10671-like	1517	96
1049	gi16198091	Drosophila melanogaster	LD30661p	184	30
1050	AAG81431	Homo sapiens	Human AFP protein sequence SEQ ID NO:380.	503	97
1050	gi6707026	Monodelphis domestica	immunoglobulin light chain kappa	108	26
1050	gi6653413	Oryctolagus cuniculus	immunoglobulin light chain VJ kappa region	102	27
1051	gi12836893	Gallus gallus	IPR328-like protein	158	29
1051	gi3093433	Homo sapiens	Chromosome 16 BAC clone CIT987SK-625P11, complete sequence.	151	29
1051	gi4558766	Homo sapiens	neuronal voltage gated calcium channel gamma-3 subunit mRNA, complete cds.	151	29
1052	gi4337100	Homo sapiens	MSH55 gene, partial cds; and CLIC1, DDAH, G6b, G6c, G5b, G6d, G6e, G6f, BAT5, G5b, CSK2B, BAT4, G4, Apo M, BAT3, BAT2, AIF-1, 1C7, LST-1, LTB, TNF, and LTA genes, complete cds.	400	100
1052	gi5304878	Homo sapiens	genes encoding RNCC protein, DDAH protein, Ly6-C protein, Ly6-D protein and immunoglobulin receptor.	400	100
1052	AAY27597	Homo sapiens	Human secreted protein encoded by gene No. 31.	400	100
1053	AAB88325	Homo sapiens	Human membrane or secretory protein clone PSEC0020.	912	99
1053	AAB53257	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:797.	859	99
1053	gi13325409	Homo sapiens	clone IMAGE:3845253, mRNA, partial cds.	774	100
1054	gi1234787	Xenopus laevis	up-regulated by thyroid hormone in tadpoles; expressed specifically in the tail and only at metamorphosis; membrane bound or extracellular protein; C-terminal basic region	917	61
1054	gi10435980	Homo sapiens	cDNA FLJ13840 fis, clone	812	62

Table 2A

142

SEQ ID	Hit ID	Species	Description	S score	% Identity
			THYRO1000783, moderately similar to <i>Xenopus laevis</i> tail-specific thyroid hormone up-regulated (gene 5) mRNA.		
1054	AAB94773	Homo sapiens	Human protein sequence SEQ ID NO:15860.	812	62
1055	gi4099139	Homo sapiens	Human P2X4 purinoreceptor gene, exons 9, 10, 11 and 12 and complete cds.	2014	100
1055	gi4099121	Homo sapiens	Human P2X4 purinoreceptor mRNA, complete cds.	2014	100
1055	AAW47066	Homo sapiens	Human brain P2X-1 receptor polypeptide.	2014	100
1056	AAE03560	Homo sapiens	Human differentially expressed kidney cDNA 22360 encoded protein.	1020	99
1056	gi15637151	Beta vulgaris	glycine decarboxylase subunit P	62	36
1056	gi5824822	Caenorhabditis elegans	Y53F4A.2	62	25
1057	gi972946	Mus musculus	ZP1 precursor	2217	67
1057	gi1113794	Mus musculus	zona pellucida	2210	67
1057	gi2804566	Rattus norvegicus	zona pellucida 1 glycoprotein	2200	67
1058	gi15779156	Homo sapiens	Similar to RIKEN cDNA 1810073N04 gene, clone MGC:15523 IMAGE:3028844, mRNA, complete cds.	1858	100
1058	gi13097045	Mus musculus	Similar to RIKEN cDNA 1810073N04 gene	1719	91
1058	gi603254	Saccharomyces cerevisiae	Yel064cp	319	27
1059	AAW03516	Homo sapiens	Prostaglandin DP receptor.	1467	100
1059	gi940379	Homo sapiens	Human DP prostanoid receptor (PTGDR) gene, 5' region and partial cds.	1467	100
1059	gi4567038	Rattus norvegicus	prostaglandin D2 receptor	1127	77
1060	gi2811122	Xenopus laevis	NaDC-2	1274	56
1060	gi1098557	Homo sapiens	Human renal sodium/dicarboxylate cotransporter (NADC1) mRNA, complete cds.	1618	55
1060	gi3168585	Rattus norvegicus	sodium-dependent dicarboxylate transporter	1614	54
1061	gi3036840	Homo sapiens	mRNA for cystinosin.	1686	88
1061	gi3036851	Homo sapiens	CTNS gene, exon 3, flanking intronic regions and joined CDS.	1686	88
1061	gi7239176	Homo sapiens	vanilloid receptor gene, partial sequence; CARKL and CTNS genes, complete cds; TIP1 gene, partial cds; P2X5b and P2X5a genes, complete cds; and HUMINAE gene, partial cds.	1686	88
1062	gi41077	Escherichia coli	cal protein precursor (aa 1-51)	63	42
1062	gi6474978	Schizosaccharomyces pombe	Amino acid permease	62	27
1062	AAB40157	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:67.	60	27

Table 2A

143

SEQ ID	Hit ID	Species	Description	S score	% Identity
1063	AAY36071	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 456.	1252	92
1063	gi15990604	Homo sapiens	RAE-1-like transcript 4 mRNA, complete cds.	1022	97
1063	AAG00501	Homo sapiens	Human secreted protein, SEQ ID NO: 4582.	533	95
1064	gi14290560	Homo sapiens	Similar to transmembrane 7 superfamily member 2, clone MGC:9286 IMAGE:3874367, mRNA, complete cds.	1548	98
1064	gi15277509	Homo sapiens	Similar to transmembrane 7 superfamily member 2, clone MGC:17157 IMAGE:4214662, mRNA, complete cds.	1548	97
1064	gi3211722	Homo sapiens	lamin B receptor homolog TM7SF2 (TM7SF2) mRNA, complete cds.	1132	77
1066	AAE06611	Homo sapiens	Human protein having hydrophobic domain, HP03696.	1552	99
1066	gi13676372	Homo sapiens	clone MGC:4595 IMAGE:3345729, mRNA, complete cds.	469	50
1066	AAY41690	Homo sapiens	Human PRO329 protein sequence.	469	50
1067	AAG72119	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1800.	1036	81
1067	gi3769616	Rattus norvegicus	olfactory receptor	887	81
1067	gi12054453	Homo sapiens	6M1-18*01 gene for olfactory receptor, cell line BM28.7.	547	42
1068	gi7106778	Homo sapiens	HSPC194	530	95
1068	AAW64547	Homo sapiens	Human stomach cancer cell clone HP10175 protein.	530	95
1068	AAY35949	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 198.	530	95
1069	gi402185	Homo sapiens	H.sapiens ALK-2 mRNA.	1572	100
1069	gi338219	Homo sapiens	Human novel serine kinase receptor mRNA, complete cds.	1572	100
1069	AAR85206	Homo sapiens	Human ALK-2.	1572	100
1070	gi4128041	Homo sapiens	claudin-9 (CLDN9) gene.	227	35
1070	AAB64401	Homo sapiens	Amino acid sequence of human intracellular signalling molecule INTRA33.	227	35
1070	gi4325296	Mus musculus	claudin-9	214	34
1071	gi1405893	Homo sapiens	H.sapiens MICA gene.	1896	93
1071	AAW60043	Homo sapiens	Human MHC class I chain-related gene A (MICA) polypeptide.	1896	93
1071	gi508492	Homo sapiens	Human MHC class I-related protein mRNA, complete cds.	1838	90
1072	gi15292437	Drosophila melanogaster	LP10272p	444	39
1072	gi4877582	Homo sapiens	lipoma HMGIC fusion partner (LHFP) mRNA, complete cds.	221	28
1072	AAY87336	Homo sapiens	Human signal peptide containing protein HSPP-113 SEQ ID NO:113.	221	28
1073	AAB58289	Homo sapiens	Lung cancer associated	1338	100

Table 2A

144

SEQ ID	Hit ID	Species	Description	S score	% Identity
			polypeptide sequence SEQ ID 627.		
1073	AAY29332	Homo sapiens	Human secreted protein clone pe584_2 protein sequence.	1338	100
1073	AAB75295	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:114.	1247	100
1074	AAB58289	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 627.	1338	100
1074	AAY29332	Homo sapiens	Human secreted protein clone pe584_2 protein sequence.	1338	100
1074	AAB75295	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:114.	1247	100
1075	AAE04780	Homo sapiens	Human vesicle trafficking protein-23 (VETRP-23) protein.	864	100
1075	AAB28629	Homo sapiens	Human B11Ag1 antigen splice isoform B11C-8.	546	39
1075	AAB28630	Homo sapiens	Human B11Ag1 antigen splice isoform B11C-9-16.	546	39
1076	gi15811373	Mus musculus	G protein coupled receptor affecting testicular descent	1269	83
1076	gi10441730	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 7 (LGR7) mRNA, complete cds.	1004	62
1076	AAY42170	Homo sapiens	Human LGR7 long form protein sequence.	1004	62
1077	gi13544043	Homo sapiens	clone IMAGE:3627317, mRNA, partial cds.	1257	52
1077	gi14249892	Homo sapiens	spinster-like protein, clone MGC:15767 IMAGE:3501826, mRNA, complete cds.	1257	52
1077	gi12003980	Homo sapiens	spinster-like protein mRNA, complete cds.	1257	52
1078	AAB85029	Homo sapiens	Protein encoded by BAP28 cDNA consisting of exons 1 to 45.	1618	68
1078	AAW54099	Homo sapiens	Homo sapiens BAP28 sequence.	1617	67
1078	gi7022341	Homo sapiens	cDNA FLJ10359 fis, clone NT2RM2001243.	1588	92
1079	gi13491841	Rattus norvegicus	gamma-glutamyltranspeptidase-like protein	209	34
1079	AAG75266	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6030.	217	100
1079	gi57806	Rattus sp.	gamma-glutamyltranspeptidase (AA 1-568)	186	33
1080	gi5262646	Homo sapiens	mRNA; cDNA DKFZp434I091 (from clone DKFZp434I091); partial cds.	2917	100
1080	gi6807820	Homo sapiens	mRNA; cDNA DKFZp434A2372 (from clone DKFZp434A2372); partial cds.	629	100
1080	gi1408182	Homo sapiens	Human LGN protein mRNA, complete cds.	282	31
1081	gi4878022	Homo sapiens	acyl-coenzyme A: cholesterol acyltransferase mRNA, complete cds.	930	98
1081	AAR53079	Homo sapiens	Acetyl coenzyme A: cholesterol acetyltransferase (ACAT).	925	98



Table 2A  
145

SEQ ID	Hit ID	Species	Description	S score	% Identity
1081	AAW38416	Homo sapiens	Human acyl-coenzyme A:cholesterol acyltransferase I.	925	98
1082	gi458938	Saccharomyces cerevisiae	Yhr186cp	1004	58
1082	gi5921144	Schizosaccharomyces pombe	mip1	2049	52
1082	gi9366720	Trypanosoma brucei	possible t16o11.22 protein.	277	45
1083	gi402187	Homo sapiens	H.sapiens ALK-3 mRNA.	1664	99
1083	AAR55368	Homo sapiens	Human Activin receptor-like kinase 3 (hALK-3).	1664	99
1083	AAR85207	Homo sapiens	Human ALK-3.	1664	99
1084	gi609354	Xenopus laevis	BMP receptor	1485	90
1084	gi2446992	Xenopus laevis	'BMP receptor'	1483	89
1084	gi3551073	Danio rerio	type I serin/threonine kinase receptor	1451	87
1085	AAW90873	Homo sapiens	Human brain-specific dysferlin protein.	1340	53
1085	gi3600028	Homo sapiens	dysferlin mRNA, complete cds.	1340	53
1085	AAW82643	Homo sapiens	Human dysferlin protein sequence SEQ ID NO:2.	1340	53
1086	gi3600028	Homo sapiens	dysferlin mRNA, complete cds.	1866	49
1086	AAW82643	Homo sapiens	Human dysferlin protein sequence SEQ ID NO:2.	1866	49
1086	AAW90868	Homo sapiens	Human dysferlin protein.	1866	49
1087	AAW92321	Homo sapiens	Human alpha-2-delta-D calcium channel subunit.	5881	99
1087	AAB62262	Homo sapiens	Human calcium channel alpha2delta subunit.	5745	99
1087	AAW92323	Homo sapiens	Human alpha-2-delta-D polypeptide from splice variant 1.	4976	99
1088	gi2104689	Mus musculus	alpha glucosidase II, alpha subunit	1796	55
1088	gi1890664	Sus scrofa	glucosidase II	1792	55
1088	gi7672977	Homo sapiens	glucosidase II alpha subunit mRNA, complete cds.	1783	55
1089	AAW01143	Homo sapiens	Secreted protein encoded by gene 9 clone HSIDY06.	238	100
1089	gi6692409	Otus longicornis	cytochrome b	64	38
1089	gi10312185	Otus watsonii	cytochrome b	61	43
1090	gi13477285	Homo sapiens	structure specific recognition protein 1, clone MGC:1608 IMAGE:3536048, mRNA, complete cds.	3683	100
1090	gi184242	Homo sapiens	Human high mobility group box (SSRP1) mRNA, complete cds.	3683	100
1090	AAR38744	Homo sapiens	Human SSRP.	3683	100
1091	gi177814	Homo sapiens	Human alpha-1-antitrypsin-related protein gene, exons 3, 4 and 5.	1925	90
1091	AAP50132	Homo sapiens	Sequence of the predominant form of human alpha-1-antitrypsin(AT).	828	59
1091	gi15990507	Homo sapiens	Similar to serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1, clone MGC:23330 IMAGE:4644658, mRNA, complete cds.	1409	66

Table 2A

146

SEQ ID	Hit ID	Species	Description	S score	% Identity
1092	AAB56819	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1397.	1054	100
1092	gi15981490	Yersinia pestis	protease	137	28
1092	gi9654995	Vibrio cholerae	protease DegS	135	29
1093	gi13543976	Homo sapiens	clone IMAGE:3603998, mRNA, partial cds.	1523	100
1093	gi15930240	Homo sapiens	Similar to CAP-binding protein complex interacting protein 2, clone MGC:9962 IMAGE:3878011, mRNA, complete cds.	1523	100
1093	AAV57946	Homo sapiens	Human transmembrane protein HTPN-70.	1128	100
1094	AAV53031	Homo sapiens	Human secreted protein clone dd426_1 protein sequence SEQ ID NO:68.	590	93
1094	AAV71062	Homo sapiens	Human membrane transport protein, MTRP-7.	158	26
1094	gi15529155	Arabidopsis thaliana	AT3g30390/T6J22_16	135	22
1095	gi4959568	Homo sapiens	nuclear pore complex interacting protein NPIP (NPIP) mRNA, complete cds.	1650	94
1095	gi2342743	Homo sapiens	Human Chromosome 16 BAC clone CIT987SK-A-589H1, complete sequence.	1627	93
1095	AAV10915	Homo sapiens	Amino acid sequence of a human secreted peptide.	760	88
1096	gi7022118	Homo sapiens	cDNA FLJ10213 fis, clone HEMBA1006474, weakly similar to 40 KD PROTEIN.	1074	99
1096	AAB92609	Homo sapiens	Human protein sequence SEQ ID NO:10874.	1074	99
1096	gi456886	Borna disease virus	p40	396	41
1097	gi38432	Homo sapiens	H.sapiens gene for mitochondrial ATP synthase c subunit (P2 form).	612	90
1097	gi285910	Homo sapiens	P2 mRNA for ATP synthase subunit c, complete cds.	612	90
1097	AAB43694	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1139.	612	90
1098	gi897827	Homo sapiens	Human iron-responsive element-binding protein/iron regulatory protein 2 (IRE-BP2/IRP2) mRNA, partial cds.	4968	99
1098	gi897581	Homo sapiens	Human iron-regulatory protein 2 (IRP2) mRNA, partial cds.	4909	99
1098	gi897583	Rattus norvegicus	iron-regulatory protein 2	4700	93
1099	gi5732908	Homo sapiens	BPAG1n3 (BPAG1) mRNA, partial cds.	75	32
1099	AAV87302	Homo sapiens	Human signal peptide containing protein HSPP-79 SEQ ID NO:79.	61	35
1099	AAV76213	Homo sapiens	Human secreted protein encoded by gene 90.	61	35
1100	gi5732908	Homo sapiens	BPAG1n3 (BPAG1) mRNA,	75	32

Table 2A

147

SEQ ID	Hit ID	Species	Description	S score	% Identity
			partial cds.		
1100	AAY87302	Homo sapiens	Human signal peptide containing protein HSPP-79 SEQ ID NO:79.	61	35
1100	AAY76213	Homo sapiens	Human secreted protein encoded by gene 90.	61	35
1101	gi5732908	Homo sapiens	BPAG1n3 (BPAG1) mRNA, partial cds.	75	32
1101	AAY87302	Homo sapiens	Human signal peptide containing protein HSPP-79 SEQ ID NO:79.	61	35
1101	AAY76213	Homo sapiens	Human secreted protein encoded by gene 90.	61	35
1102	AAY86234	Homo sapiens	Human secreted protein HNTNC20, SEQ ID NO:149.	88	31
1102	gi5430769	Arabidopsis thaliana	Similar to somatic embryogenesis receptor-like kinase	88	32
1102	AAB24074	Homo sapiens	Human PRO1153 protein sequence SEQ ID NO:49.	79	22
1103	gi13447199	Homo sapiens	sphingosine-1-phosphate phosphatase mRNA, complete cds.	1931	98
1103	gi9623190	Mus musculus	sphingosine-1-phosphate phosphohydrolase	1692	83
1103	gi15778670	Mus musculus	sphingosine-1-phosphate phosphatase	1692	83
1104	gi12052824	Homo sapiens	mRNA; cDNA DKFZp564H1562 (from clone DKFZp564H1562); complete cds.	1544	100
1104	gi5326797	Homo sapiens	junctional adhesion molecule (JAM1) mRNA, complete cds.	1544	100
1104	gi5731339	Homo sapiens	junctional adhesion molecule-1 mRNA, complete cds.	1544	100
1105	gi296636	Homo sapiens	Human apoC-II gene for preproapolipoprotein C-II.	506	100
1105	gi757915	Homo sapiens	Human mRNA for lipoprotein apoCII.	506	100
1105	gi178836	Homo sapiens	APOC2 gene, complete sequence; and apolipoprotein C-II (APOC2) gene, complete cds.	506	100
1106	gi13097159	Homo sapiens	tumor protein, translationally-controlled 1, clone MGC:5308 IMAGE:2899964, mRNA, complete cds.	794	97
1106	gi7573519	Homo sapiens	TPT1 gene for translationally controlled tumor protein (TCTP), exons 1-6.	794	97
1106	gi37496	Homo sapiens	Human mRNA for translationally controlled tumor protein.	794	97
1107	gi12082725	Mus musculus	B cell phosphoinositide 3-kinase adaptor	3523	84
1107	gi12082723	Gallus gallus	B cell phosphoinositide 3-kinase adaptor	2821	69
1107	AAB43816	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1261.	1257	98
1108	gi10177622	Arabidopsis thaliana	gene_id:K6M13.11~	201	39
1108	gi10437414	Homo sapiens	cDNA: FLJ21330 fis, clone COL02466.	165	34

Table 2A

148

SEQ ID	Hit ID	Species	Description	S score	% Identity
1108	gi499199	Schizosaccharom yces pombe	uvi22	155	33
1109	gi13436446	Homo sapiens	myosin regulatory light chain, clone MGC:4405 IMAGE:2906108, mRNA, complete cds.	881	99
1109	gi829623	Homo sapiens	Human myosin regulatory light chain mRNA, complete cds.	881	99
1109	gi15076511	Homo sapiens	MLC-2 mRNA for nonmuscle myosin light chain 2, complete cds.	881	99
1110	gi5305502	Mus musculus	phospholemman precursor	153	45
1110	gi1916012	Rattus norvegicus	phospholemman chloride channel	142	53
1110	gi1916010	Homo sapiens	Human phospholemman chloride channel mRNA, complete cds.	133	47
1111	gi13272522	Homo sapiens	transcription factor NYD-sp10 mRNA, complete cds.	1344	90
1111	gi14278918	Homo sapiens	mRNA for transcription factor RFX4, complete cds.	1166	82
1111	gi583352	synthetic construct	does not include the start ot stop codon	162	29
1112	AAB47296	Homo sapiens	PRO4401 polypeptide.	1062	58
1112	AAY22496	Homo sapiens	Human secreted protein sequence clone cn621_8.	1062	58
1112	gi14042441	Homo sapiens	cDNA FLJ14724 fis, clone NT2RP3001716.	400	43
1113	gi15341863	Homo sapiens	Similar to RIKEN cDNA 2900052H21 gene, clone MGC:21625 IMAGE:4214683, mRNA, complete cds.	758	98
1113	AAY33297	Homo sapiens	Human membrane spanning protein MSP-4.	758	98
1113	AAB61149	Homo sapiens	Human NOV18 protein.	758	98
1114	gi11125139	Homo sapiens	Novel human gene mapping to chomosome 22.	476	89
1114	AAY94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	476	89
1114	gi602584	Methanosarcina mazei	cytochrome b	75	33
1115	AAG72267	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1948.	1281	100
1115	AAG72407	Homo sapiens	Human OR-like polypeptide query sequence, SEQ ID NO: 2088.	1281	100
1115	AAG72270	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1951.	997	73
1116	gi12002782	Homo sapiens	olfactory receptor-like protein JCG2 (JCG2) mRNA, partial cds.	1538	100
1116	gi12002784	Homo sapiens	olfactory receptor-like protein JCG2 (JCG2) gene, complete cds.	1538	100
1116	AAE04555	Homo sapiens	Human G-protein coupled receptor-11 (GCREC-11) protein.	1538	100
1117	gi5802817	Homo sapiens	endogenous retrovirus HERV-K104 long terminal repeat, complete sequence; and Gag protein (gag) and envelope protein (env) genes, complete cds.	479	77

Table 2A  
149

SEQ ID	Hit ID	Species	Description	S score	% Identity
1117	gi1469243	Human endogenous retrovirus K	pol/env	466	77
1117	gi3150438	Human endogenous retrovirus K	pol-env	466	77
1118	AAG89341	Homo sapiens	Human secreted protein, SEQ ID NO: 461.	501	99
1118	gi6651037	Mus musculus domesticus	similar to RNA binding protein	411	96
1118	AAG02095	Homo sapiens	Human secreted protein, SEQ ID NO: 6176.	167	55
1119	AAB20155	Homo sapiens	Secreted protein SECP1.	3983	51
1119	gi3080663	Homo sapiens	PAC clone RP5-1168D11 from 7p21-p22, complete sequence.	1408	47
1119	gi2897863	Homo sapiens	BAC clone GS1-164B5 from 7p21-p22, complete sequence.	1340	50
1120	gi32329	Homo sapiens	Human HMG-17 gene for non-histone chromosomal protein HMG-17.	429	94
1120	gi306864	Homo sapiens	Human non-histone chromosomal protein HMG-17 mRNA, complete cds.	429	94
1120	AAB28199	Homo sapiens	Human HMG-17 non histone chromosomal protein.	429	94
1121	gi13905022	Homo sapiens	Similar to interferon induced transmembrane protein 3 (1-8U), clone MGC:5225 IMAGE:2986145, mRNA, complete cds.	444	69
1121	gi14250038	Homo sapiens	Similar to interferon induced transmembrane protein 3 (1-8U), clone MGC:14565 IMAGE:4075453, mRNA, complete cds.	436	68
1121	gi23398	Homo sapiens	Human 1-8U gene from interferon-inducible gene family.	435	67
1122	gi7019933	Homo sapiens	cDNA FLJ20071 fis, clone COL01887.	2163	100
1122	AAB36618	Homo sapiens	Human FLEXHT-40 protein sequence SEQ ID NO:40.	1051	100
1122	AAW88957	Homo sapiens	Polypeptide fragment encoded by gene 128.	902	100
1123	AAB60112	Homo sapiens	Human transport protein TPPT-32.	775	100
1123	gi11558029	Homo sapiens	boct gene for organic cation transporter.	382	48
1123	gi9663117	Homo sapiens	mRNA for organic cation transporter.	382	48
1124	AAR28120	Homo sapiens	NKG2 transmembrane protein-D.	727	95
1124	gi2980865	Homo sapiens	NKG2D gene, exons 2-5 and joined mRNA and CDS.	724	94
1124	gi35063	Homo sapiens	Human mRNA for NKG2-D gene.	724	94
1125	gi7767239	Homo sapiens	nectin-like protein 2 (NECL2) mRNA, complete cds.	612	39
1125	gi4519602	Homo sapiens	IGSF4 gene, exon 10 and complete cds.	609	38
1125	AAAY45092	Homo sapiens	Human lymphoid derived dendritic	609	38

Table 2A

150

SEQ ID	Hit ID	Species	Description	S score	% Identity
			cell adhesion molecule.		
1126	gi7020365	Homo sapiens	cDNA FLJ20336 fis, clone HEP11722.	4316	99
1126	gi10435830	Homo sapiens	cDNA FLJ13727 fis, clone PLACE3000103.	3079	99
1126	AAB94738	Homo sapiens	Human protein sequence SEQ ID NO:15776.	3079	99
1127	AAB75594	Homo sapiens	Human secreted protein sequence encoded by gene 37 SEQ ID NO:148.	678	99
1127	AAB75542	Homo sapiens	Human secreted protein sequence encoded by gene 37 SEQ ID NO:96.	294	100
1127	gi1864011	Homo sapiens	mRNA for SHPS-1, complete cds.	261	43
1128	gi7020372	Homo sapiens	cDNA FLJ20340 fis, clone HEP12374.	1692	99
1128	gi4098525	Prochlorothrix hollandica	CytM	80	31
1128	gi324932	Influenza A virus	PA polymerase	67	38
1129	gi7023403	Homo sapiens	cDNA FLJ11006 fis, clone PLACE1003045.	499	59
1129	AAB93412	Homo sapiens	Human protein sequence SEQ ID NO:12616.	499	59
1129	gi13542919	Mus musculus	Similar to mucolipin 1	432	61
1130	gi15488920	Homo sapiens	Similar to RIKEN cDNA 2010107G23 gene, clone MGC:9596 IMAGE:3896656, mRNA, complete cds.	107	42
1130	AAW74777	Homo sapiens	Human secreted protein encoded by gene 48 clone HFCAI74.	74	40
1130	gi1304441	Pseudorabies virus	Rsp40	69	32
1131	gi10119918	Homo sapiens	brain otoferlin short isoform (OTOF) mRNA, complete cds.	1315	49
1131	gi10119916	Homo sapiens	brain otoferlin long isoform (OTOF) mRNA, complete cds.	1315	49
1131	gi4588470	Homo sapiens	otoferlin (OTOF) mRNA, complete cds.	2214	43
1132	gi1006665	Homo sapiens	H.sapiens mRNA for transcript associated with monocyte to macrophage differentiation.	442	98
1132	gi15155898	Agrobacterium tumefaciens	AGR_C_1653p	167	31
1132	gi15023850	Clostridium acetobutylicum	Predicted membrane protein, hemolysin III homolog	117	44
1133	AAG71803	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1484.	1494	92
1133	AAG71805	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1486.	1205	92
1133	AAG71807	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1488.	1178	70
1134	AAV70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	609	91
1134	AAV83992_aal	Homo sapiens	Nucleic acid encoding a protein with water channel activity.	608	92
1134	gi2317274	Homo sapiens	mRNA for aquaporin adipose,	608	92

Table 2A  
151

SEQ ID	Hit ID	Species	Description	S score	% Identity
			complete cds.		
1135	gi3319326	Homo sapiens	protein associated with Myc mRNA, complete cds.	111	33
1136	gi2463632	Homo sapiens	monocarboxylate transporter homologue MCT6 mRNA, complete cds.	2574	97
1136	gi10880482	Mus musculus	monocarboxylate transporter 4	393	39
1136	gi2463634	Homo sapiens	monocarboxylate transporter (MCT3) mRNA, complete cds.	394	40
1137	gi13528675	Homo sapiens	ATPase, H <sup>+</sup> transporting, lysosomal (vacuolar proton pump) 16kD, clone MGC:3723 IMAGE:3618755, mRNA, complete cds.	705	94
1137	gi13938484	Homo sapiens	ATPase, H <sup>+</sup> transporting, lysosomal (vacuolar proton pump) 16kD, clone MGC:16271 IMAGE:3831016, mRNA, complete cds.	705	94
1137	gi14043553	Homo sapiens	ATPase, H <sup>+</sup> transporting, lysosomal (vacuolar proton pump) 16kD, clone MGC:12873 IMAGE:4127653, mRNA, complete cds.	705	94
1138	gi15080314	Homo sapiens	Similar to RIKEN cDNA 0610010D20 gene, clone MGC:20590 IMAGE:4310241, mRNA, complete cds.	514	100
1138	gi10580053	Halobacterium sp. NRC-1	dihydrodipicolinate synthase; DapA	379	33
1138	gi1590977	Methanococcus jannaschii	dihydrodipicolinate synthase (dapA)	336	29
1139	AAE06614	Homo sapiens	Human protein having hydrophobic domain, HP03974.	1394	100
1139	gi520469	Oryctolagus cuniculus	597 aa protein related to Na/glucose cotransporters	1231	85
1139	gi338055	Homo sapiens	Human Na <sup>+</sup> /glucose cotransporter 1 mRNA, complete cds.	705	57
1140	gi6708478	Mus musculus	formin-like protein	1571	66
1140	gi4101720	Mus musculus	lymphocyte specific formin related protein	1543	65
1140	gi1914849	Mus musculus	WW domain binding protein 3; WBP3	299	54
1142	gi12052738	Homo sapiens	mRNA; cDNA DKFZp564H1322 (from clone DKFZp564H1322); complete cds.	1755	96
1142	gi10434632	Homo sapiens	cDNA FLJ12886 fis, clone NT2RP2004041, weakly similar to SYNAPSINS IA AND IB.	1755	96
1142	AAB94358	Homo sapiens	Human protein sequence SEQ ID NO:14883.	1755	96
1143	AAW54370	Homo sapiens	G-protein coupled receptor HLTEX11.	1815	100
1143	AAB64854	Homo sapiens	Human secreted protein sequence encoded by gene 36 SEQ ID NO:140.	1792	100
1143	AAW70504	Homo sapiens	Leukocyte seven times membrane-	821	46

Table 2A

152

SEQ ID	Hit ID	Species	Description	S score	% Identity
			penetrating type receptor protein JEG18.		
1144	gi15278128	Mus musculus	chemokine-like factor 2 variant 2	412	49
1144	AAB51648	Homo sapiens	Human secreted protein sequence encoded by gene 29 SEQ ID NO:88.	410	100
1144	AAE03929	Homo sapiens	Human gene 32 encoded secreted protein HTLIF12, SEQ ID NO:92.	410	100
1146	gi13477335	Homo sapiens	vitamin A responsive; cytoskeleton related, clone MGC:1917 IMAGE:3510436, mRNA, complete cds.	777	95
1146	gi3746652	Homo sapiens	JWA protein mRNA, complete cds.	777	95
1146	gi6563260	Homo sapiens	jmx protein mRNA, complete cds.	777	95
1147	gi2970431	Florometra serratissima	NADH dehydrogenase subunit 4	94	31
1147	gi15042530	Chilo iridescent virus	450L	70	24
1147	AAV87197	Homo sapiens	Human secreted protein sequence SEQ ID NO:236.	90	27
1148	AAB93562	Homo sapiens	Human protein sequence SEQ ID NO:12957.	2402	100
1148	gi7023538	Homo sapiens	cDNA FLJ11091 fis, clone PLACE1005313.	860	100
1148	AAB93489	Homo sapiens	Human protein sequence SEQ ID NO:12790.	860	100
1150	gi10438431	Homo sapiens	cDNA: FLJ22155 fis, clone HRC00205.	1995	100
1150	gi10437336	Homo sapiens	cDNA: FLJ21267 fis, clone COL01717.	1776	99
1150	gi7020065	Homo sapiens	cDNA FLJ20152 fis, clone COL08515.	705	100
1151	gi12654159	Homo sapiens	interferon induced transmembrane protein 1 (9-27), clone MGC:5195 IMAGE:3464598, mRNA, complete cds.	569	93
1151	gi1177476	Homo sapiens	H.sapiens mRNA for interferon-induced 17kDa membrane protein.	569	93
1151	gi177802	Homo sapiens	Human interferon-inducible protein 9-27 mRNA, complete cds.	563	92
1152	AAG72230	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1911.	1615	100
1152	AAG72382	Homo sapiens	Human OR-like polypeptide query sequence, SEQ ID NO: 2063.	1615	100
1152	gi15293613	Homo sapiens	clone OR5C1 olfactory receptor gene, partial cds.	1097	100
1153	gi784997	Homo sapiens	H.sapiens mRNA for tumour suppressor protein, HUGL.	5025	95
1153	gi1944491	Homo sapiens	Human LLGL mRNA, complete cds.	4797	91
1153	gi854124	Homo sapiens	H.sapiens mRNA for human giant larvae homolog.	2837	58
1154	AAB95830	Homo sapiens	Human protein sequence SEQ ID NO:18850.	219	72
1154	gi7959889	Homo sapiens	PRO2221	137	49



Table 2A  
153

SEQ ID	Hit ID	Species	Description	S score	% Identity
1154	gi2072969	Homo sapiens	Human L1 element L1.24 p40 gene, complete cds.	133	48
1155	gi15277644	Homo sapiens	amino acid transporter (SLC7A10) gene, exon 11 and complete cds.	2487	100
1155	gi9309293	Homo sapiens	hasc-1 mRNA for asc-type amino acid transporter 1, complete cds.	2487	100
1155	gi7415938	Mus musculus	asc1	2329	91
1156	gi6760373	Homo sapiens	ODZ3 (ODZ3) mRNA, partial cds.	2323	100
1156	gi4760780	Mus musculus	Ten-m3	2248	96
1156	gi6010049	Gallus gallus	teneurin-2 protein	878	62
1157	gi14286298	Homo sapiens	clone MGC:3593 IMAGE:2963628, mRNA, complete cds.	630	99
1157	gi4877285	Homo sapiens	mRNA for prenylated Rab acceptor 1.	630	99
1157	gi6563192	Homo sapiens	prenylated rab acceptor 1 mRNA, complete cds.	630	99
1158	gi1780976	Human endogenous retrovirus K	protease	915	58
1158	gi5802824	Homo sapiens	endogenous retrovirus HERV-K109, complete sequence.	909	59
1158	gi9558703	Homo sapiens	tandemly repeated human endogenous retrovirus HERV-K (HML-2.HOM), complete sequence.	905	59
1159	gi13111941	Homo sapiens	vesicle-associated soluble NSF attachment protein receptor (v-SNARE; homolog of S. cerevisiae VTI1), clone MGC:3767 IMAGE:2958320, mRNA, complete cds.	804	91
1159	gi3861488	Homo sapiens	vesicle soluble NSF attachment protein receptor VTI2 mRNA, complete cds.	804	91
1159	AAY73339	Homo sapiens	HTRM clone 2056042 protein sequence.	804	91
1160	gi1922891	Mus musculus	alpha 3B chain of laminin-5	10355	75
1160	gi5777581	Homo sapiens	H.sapiens mRNA for laminin-5, alpha3b chain.	9398	99
1160	gi551597	Homo sapiens	laminin-related protein (LamA3) mRNA, complete cds.	8690	100
1161	AAD05056_aal	Homo sapiens	HUMA- Human secreted protein-encoding gene 4 cDNA clone HKAAV61, SEQ ID NO:14.	1524	83
1161	AAE01167	Homo sapiens	HUMA- Human gene 4 encoded secreted protein HKAAV61, SEQ ID NO:68.	1523	82
1161	AAG67516	Homo sapiens	SMIK Amino acid sequence of a human secreted polypeptide.	1523	82
1162	AAM42034	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 6965.	900	99
1162	AAM40248	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 3393.	821	100
1162	gi491284	synthetic construct	IFN-pseudo-omega 2	800	98
1163	AAI70234	Homo sapiens	AMGE- Human interleukin-1	819	100

Table 2A  
154

SEQ ID	Hit ID	Species	Description	S score	% Identity
	aal		receptor antagonist related protein cDNA.		
1163	AAD11148_aal	Homo sapiens	IMMV Human FIL-1 theta full-length DNA.	819	100
1163	AAA89175_aal	Homo sapiens	HYSE- Human interleukin-1 Hy2 cDNA.	819	100
1164	gi340215	Homo sapiens	Human vascular endothelial growth factor gene, exon 8.	1056	97
1164	gi340301	Homo sapiens	Human vascular permeability factor mRNA, complete cds.	1056	97
1164	AAR91077	Homo sapiens	PRIZ- Human vascular endothelial growth factor-189, VEGF-189.	1056	97
1166	gi1321816	Gorilla gorilla	interleukin-8 receptor type B	602	90
1166	gi1109691	Homo sapiens	Human interleukin-8 receptor type B (IL8RB) mRNA, complete cds.	599	88
1166	gi186378	Homo sapiens	Human interleukin 8 receptor B mRNA, complete cds.	599	88
1167	gi1160967	Homo sapiens	palmitoyl-protein thioesterase gene, complete cds.	1285	100
1167	gi1314355	Homo sapiens	Human palmitoyl protein thioesterase mRNA, complete cds.	1285	100
1167	gi14250054	Homo sapiens	palmitoyl-protein thioesterase 1 (ceroid-lipofuscinosis, neuronal 1, infantile), clone MGC:14590 IMAGE:4249991, mRNA, complete cds.	1285	100
1168	gi177814	Homo sapiens	Human alpha-1-antitrypsin-related protein gene, exons 3, 4 and 5.	1956	90
1168	AAP50132	Homo sapiens	ZYMO- Sequence of the predominant form of human alpha-1-antitrypsin(AT).	1009	72
1168	gi15990507	Homo sapiens	Similar to serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1, clone MGC:23330 IMAGE:4644658, mRNA, complete cds.	1450	69
1169	AAO12931	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 26823.	346	100
1169	AAO02697	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 16589.	143	66
1169	AAO08307	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 22199.	137	80
1170	AAR15222	Homo sapiens	TEXA Chronic myelogenous leukaemia-derived myeloid-related protein.	635	100
1170	gi181527	Homo sapiens	Human neutrophil peptide (defensin) 1 mRNA, complete cds.	493	100
1170	gi181529	Homo sapiens	Human defensin 1 protein mRNA, complete cds.	493	100
1171	gi1001697	Synechocystis sp. PCC 6803	sensory transduction histidine kinase	67	37
1171	gi2353225	Acromyrmex octospinosus	cytochrome oxidase I	54	40
1171	AAG02950	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 7031.	54	32
1172	gi4884282	Homo sapiens	mRNA; cDNA DKFZp566K0524	1998	99

Table 2A  
155

SEQ ID	Hit ID	Species	Description	S score	% Identity
			(from clone DKFZp566K0524); partial cds.		
1172	gi2665458	Mus musculus	protein-tyrosine-phosphatase	1363	63
1172	gi452194	Homo sapiens	Human mRNA for protein tyrosine phosphatase (PTP-BAS, type 3), complete cds.	618	48
1173	AAY48226	Homo sapiens	META- Human prostate cancer-associated protein 12.	956	96
1173	AAM25683	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1198.	956	96
1173	AAY99342	Homo sapiens	GETH Human PRO1018 (UNQ501) amino acid sequence SEQ ID NO:8.	950	94
1174	gi178018	Homo sapiens	Human activation (Act-2) mRNA, complete cds.	125	84
1174	gi34218	Homo sapiens	Human LAG-1 mRNA.	125	84
1174	gi533213	Homo sapiens	secreted T cell protein (H400; SIS-gamma) mRNA, complete cds.	125	84
1175	gi178018	Homo sapiens	Human activation (Act-2) mRNA, complete cds.	123	92
1175	gi34218	Homo sapiens	Human LAG-1 mRNA.	123	92
1175	gi533213	Homo sapiens	secreted T cell protein (H400; SIS-gamma) mRNA, complete cds.	123	92
1176	AAG03315	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 7396.	314	100
1176	gi965069	Serratia marcescens	phage-holin analog protein	71	40
1176	gi16415877	Octopus salutii	cytochrome oxidase subunit III	74	28
1177	gi178836	Homo sapiens	APOC2 gene, complete sequence; and apolipoprotein C-II (APOC2) gene, complete cds.	453	89
1177	gi296636	Homo sapiens	Human apoC-II gene for preproapolipoprotein C-II.	453	89
1177	gi757915	Homo sapiens	Human mRNA for lipoprotein apoCII.	453	89
1178	AAO07986	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 21878.	77	34
1179	AAB60502	Homo sapiens	INCY- Human cell cycle and proliferation protein CCYPR-50, SEQ ID NO:50.	1205	100
1179	AAB12144	Homo sapiens	PROT- Hydrophobic domain protein isolated from WERI-RB cells.	1205	100
1179	AAG81333	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:184.	687	99
1180	AAW67880	Homo sapiens	HUMA- Human secreted protein encoded by gene 74 clone HNTAC73.	378	97
1180	gi9949887	Pseudomonas aeruginosa	probable transcriptional regulator	65	41
1180	gi3130050	Schizosaccharom yces pombe	haloacid dehalogenase-like hydrolase	62	32
1181	AAG01183	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 5264.	278	94
1181	AAO00621	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 14513.	84	47

Table 2A

156

SEQ ID	Hit ID	Species	Description	S score	% Identity
1181	AAO02188	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 16080.	78	40
1182	ABB12063	Homo sapiens	HYSE- Human secreted protein homologue, SEQ ID NO:2433.	326	100
1182	AAE06730	Homo sapiens	SMIK Human CASB765 protein.	200	100
1182	AAW75098	Homo sapiens	HUMA- Human secreted protein encoded by gene 42 clone HSXBI25.	143	81
1183	gi13278924	Homo sapiens	neural proliferation, differentiation and control, 1, clone MGC:4597 IMAGE:3347743, mRNA, complete cds.	748	98
1183	gi8515886	Homo sapiens	NPDC1-like protein (NPDC1) mRNA, complete cds.	748	98
1183	AAB43904	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1349.	748	98
1184	gi13128925	Homo sapiens	ULBP2 protein mRNA, complete cds.	1025	90
1184	gi14530663	Homo sapiens	mRNA for ALCAN-alpha, complete cds.	1025	90
1184	AAV15238	Homo sapiens	IMMV ULBP-2 amino acid sequence.	1025	90
1185	gi4164134	Homo sapiens	cosmid clone U169D2 from Xp22.1-22.2, complete sequence.	76	36
1185	AAU22866	Homo sapiens	HUMA- Human prostate cancer antigen, Seq ID No 385.	60	35
1185	AAM96178	Homo sapiens	HUMA- Human reproductive system related antigen SEQ ID NO: 4836.	60	35
1186	gi7770223	Homo sapiens	PRO2714	282	84
1186	ABB15615	Homo sapiens	HUMA- Human nervous system related polypeptide SEQ ID NO 4272.	63	48
1186	AAO07531	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 21423.	67	46
1187	AAE05962	Homo sapiens	HYSE- Human phospholipase-related protein.	2521	99
1187	gi3811347	Homo sapiens	cytosolic phospholipase A2 beta (cPLA2 beta) mRNA, complete cds.	1209	44
1187	gi4886978	Homo sapiens	cytosolic phospholipase A2 beta mRNA, complete cds.	1209	44
1188	AAO01938	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 15830.	93	36
1188	gi2992470	Mus sp.	mitochondrial capsule selenoprotein; MCS	95	31
1188	gi14717800	Mus musculus	seleno-protein	95	31
1189	gi186600	Homo sapiens	Human inter-alpha-trypsin inhibitor light chain (ITI) gene, exons 7-10.	1461	93
1189	gi24479	Homo sapiens	Human mRNA for alpha-1-microglobulin and HI-30.	1461	93
1189	gi32047	Homo sapiens	Human mRNA for protein HC (alpha-1-microglobulin).	1461	93
1190	gi673422	Homo sapiens	H.sapiens mRNA fragment for T-cell receptor alpha chain.	1322	93

Table 2A  
157

SEQ ID	Hit ID	Species	Description	S score	% Identity
1190	AAU69943	Homo sapiens	CORI- Human T cell receptor Va chain of clone 4E5 for prostate protein P501S.	1097	77
1190	AAM01298	Homo sapiens	CORI- P501S-specific T cell clone 4E5 Va chain T cell receptor amino acid.	1097	77
1191	gi673422	Homo sapiens	H.sapiens mRNA fragment for T-cell receptor alpha chain.	859	95
1191	gi623119	Macaca mulatta	T-cell receptor alpha	605	86
1191	AAU69943	Homo sapiens	CORI- Human T cell receptor Va chain of clone 4E5 for prostate protein P501S.	594	65
1192	gi13097588	Homo sapiens	clone MGC:10745 IMAGE:2820343, mRNA, complete cds.	201	100
1192	gi53861	Mus musculus	Q300 protein (AA 1-77)	72	38
1192	AAO02105	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 15997.	50	56
1193	AAB08894	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 4 SEQ ID NO:51.	208	57
1193	gi15088679	Homo sapiens	cysteine and tyrosine-rich protein 1 (CYYR1) mRNA, complete cds.	59	28
1193	AAY87233	Homo sapiens	INCY- Human signal peptide containing protein HSPP-10 SEQ ID NO:10.	59	28
1194	AAG03963	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 8044.	417	80
1194	ABB10168	Homo sapiens	HUMA- Human cDNA SEQ ID NO: 476.	289	100
1194	ABB10412	Homo sapiens	HUMA- Human cDNA SEQ ID NO: 720.	289	100
1195	gi9758052	Arabidopsis thaliana		64	46
1195	gi6815537	Human immunodeficiency virus type 1	gag protein	47	62
1195	gi14269033	Sus scrofa	interferon beta-1	42	47
1196	gi7582276	Homo sapiens	BM-002	328	100
1196	AAM78626	Homo sapiens	HYSE- Human protein SEQ ID NO 1288.	328	100
1196	AAM79610	Homo sapiens	HYSE- Human protein SEQ ID NO 3256.	328	100
1197	gi1181885	Mus musculus	patched	209	62
1197	AAV21590_aal	Homo sapiens	STRD Human patched (ptc) protein encoding cDNA.	215	63
1197	gi1335864	Homo sapiens	Human patched homolog (PTC) mRNA, complete cds.	214	63
1198	AAB25674	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 10 SEQ ID NO:63.	646	84
1198	gi14603247	Homo sapiens	Similar to RIKEN cDNA 5730409G15 gene, clone MGC:19636 IMAGE:2822323, mRNA, complete cds.	420	94
1198	AAB36613	Homo sapiens	INCY- Human FLEXHT-35 protein sequence SEQ ID NO:35.	420	94

Table 2A

158

SEQ ID	Hit ID	Species	Description	S score	% Identity
1200	AAU12292	Homo sapiens	GETH Human PRO6027 polypeptide sequence.	990	98
1200	AAU27673	Homo sapiens	ZYMO Human protein AFP235412.	987	99
1200	gi13477121	Homo sapiens	clone IMAGE:3636082, mRNA, partial cds.	291	95
1201	AAB43928	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1373.	216	58
1201	gi13325337	Homo sapiens	clone MGC:10520 IMAGE:3938462, mRNA, complete cds.	219	48
1201	AAB21040	Homo sapiens	INCY- Human nucleic acid-binding protein, NuABP-44.	219	48
1202	AAB43928	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1373.	223	55
1202	gi16550327	Homo sapiens	cDNA FLJ31005 fis, clone HLUNG2000068, weakly similar to ZINC FINGER PROTEIN 157.	215	67
1202	gi16552980	Homo sapiens	cDNA FLJ32768 fis, clone TESTI2001879, weakly similar to ZINC FINGER PROTEIN 157.	215	67
1203	gi4322936	Homo sapiens	HPK/GCK-like kinase HGK mRNA, complete cds.	120	85
1203	gi4262362	Caenorhabditis elegans	alternatively spliced serine/threonine protein kinase MIG-15	119	81
1203	AAB50058	Homo sapiens	IMMV SS-4694.	117	81
1204	gi1754714	Xenopus laevis	oviductin	378	40
1204	gi15277254	Bufo japonicus	oviductin	351	36
1204	gi2981641	Xenopus laevis	polyprotein	339	46
1205	gi15214578	Homo sapiens	clone MGC:9135 IMAGE:3865141, mRNA, complete cds.	451	76
1205	AAW67901	Homo sapiens	HUMA- Human secreted protein encoded by gene 36 clone HODCL36.	451	76
1205	gi1946205	Homo sapiens	H.sapiens mRNA for CHD5 protein.	445	75
1206	AAG01971	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6052.	314	100
1206	gi4200340	Lycopersicon esculentum	P69D protein	83	43
1206	gi3183989	Lycopersicon esculentum	P69E protein	82	43
1207	gi14043211	Homo sapiens	Similar to RIKEN cDNA 4931428F04 gene, clone IMAGE:3346497, mRNA, partial cds.	878	83
1207	gi9711595	Xanthomonas oryzae pv. oryzae	HpaA	71	24
1207	AAO10768	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 24660.	72	34
1208	AAAY91512	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 62 SEQ	606	100

Table 2A  
159

SEQ ID	Hit ID	Species	Description	S score	% Identity
			ID NO:185.		
1208	AAY91653	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 62 SEQ ID NO:326.	606	100
1208	AAY71277	Homo sapiens	ZYMO Human Zlipo3 protein.	606	100
1209	AAY27648	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 82.	322	98
1209	gi7959897	Homo sapiens	PRO2379	72	39
1209	AAO03791	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 17683.	61	68
1210	gi17431247	Ralstonia solanacearum	HYPOTHETICAL PROTEIN	70	38
1211	AAB08765	Homo sapiens	INCY- A human leukocyte and blood related protein (LBAP).	339	62
1211	AAB74718	Homo sapiens	INCY- Human membrane associated protein MEMAP-24.	314	66
1211	gi2587024	Homo sapiens	mRNA containing human endogenous retrovirus H and human endogenous retrovirus E sequences.	259	67
1212	gi10440139	Homo sapiens	cDNA: FLJ23447 fis, clone HSI03346.	1339	95
1212	AAY13396	Homo sapiens	GETH Amino acid sequence of protein PRO332.	872	48
1212	AAB33425	Homo sapiens	GETH Human PRO332 protein UNQ293 SEQ ID NO:57.	872	48
1213	AAG66547	Homo sapiens	HYSE- Human secreted metallocarboxypeptidase-like polypeptide.	1551	99
1213	AAG66565	Homo sapiens	HYSE- Human secreted metallocarboxypeptidase-like variant polypeptide.	1548	98
1213	AAB74682	Homo sapiens	INCY- Human protease and protease inhibitor PPIM-15.	1482	98
1214	gi15528833	Homo sapiens	Fc receptor-like protein 2 (FCRH2) mRNA, complete cds.	528	100
1214	AAB85464	Homo sapiens	HYSE- Human immunoglobulin domain-containing polypeptide.	528	100
1214	AAB82317	Homo sapiens	UYCO Human immunoglobulin receptor IRTA4 protein.	528	100
1215	AAU27663	Homo sapiens	ZYMO Human protein AFP285042.	555	100
1215	AAE06599	Homo sapiens	SAGA Human protein having hydrophobic domain, HP10782.	510	100
1215	gi15558917	Magnaporthe grisea	similarity to Lec35 protein	169	30
1216	gi10439008	Homo sapiens	cDNA: FLJ22573 fis, clone HSI02387.	682	99
1216	AAM87876	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:15469.	169	70
1216	gi1616747	Tupaia glis belangeri	GnRH preprohormone second form	68	33
1217	gi10439008	Homo sapiens	cDNA: FLJ22573 fis, clone HSI02387.	529	92
1217	AAM87876	Homo sapiens	HUMA- Human	109	71

Table 2A

160

SEQ ID	Hit ID	Species	Description	S score	% Identity
			immune/haematopoietic antigen SEQ ID NO:15469.		
1217	AAM87620	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:15213.	70	35
1218	AAM60951	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 33056.	58	36
1218	AAM73644	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 33950.	58	36
1218	AAO00109	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 14001.	58	45
1219	gi16950537	Hop mosaic virus	cysteine-rich nucleic acid binding protein	41	47
1219	AAY19474	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	43	43
1219	AAO03071	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 16963.	45	31
1220	gi1171589	Plasmodium falciparum	frameshift	76	37
1220	gi4512010	Escherichia coli	OrfY	66	50
1220	gi1870470	Mus musculus	anti-DNA immunoglobulin light chain IgG	46	37
1221	gi13274524	Homo sapiens	complement-c1q tumor necrosis factor-related protein (CTRP7) mRNA, complete cds.	1451	94
1221	AAB50371	Homo sapiens	ZYMO Human ZACRP7.	1451	94
1221	gi13274518	Homo sapiens	complement-c1q tumor necrosis factor-related protein (CTRP2) mRNA, complete cds.	831	61
1222	AAO03899	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 17791.	67	50
1222	AAG73465	Homo sapiens	HUMA- Human gene 12-encoded secreted protein fragment, SEQ ID NO:240.	75	31
1222	AAM85406	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:12999.	60	40
1223	gi8850245	Homo sapiens	activated p21cdc42Hs kinase (ACK1) mRNA, complete cds.	5605	100
1223	gi2921447	Mus musculus	non-receptor protein tyrosine kinase Ack	4238	79
1223	gi2078388	Bos taurus	Cdc42-associated tyrosine kinase ACK-2	3493	90
1224	AAB84696	Homo sapiens	ZYMO Amino acid sequence of a human zkun10 polypeptide.	358	35
1224	gi211622	Gallus gallus	alpha-3 collagen type VI	308	33
1224	AAM42089	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 7020.	276	32
1225	AAB66065	Homo sapiens	MILL- Human TANGO 294.	2113	99
1225	AAB66067	Homo sapiens	MILL- Human TANGO 294 mature protein.	2015	99
1225	gi434306	Homo sapiens	H.sapiens mRNA for lysosomal acid lipase.	1290	60
1226	AAM06483	Homo sapiens	HYSE- Human foetal protein, SEQ	282	98



Table 2A

161

SEQ ID	Hit ID	Species	Description	S score	% Identity
			ID NO: 214.		
1226	gi3694984	Pimpinella brachycarpa	metallothionein-1-like protein	57	30
1226	AAU22415	Homo sapiens	HUMA- Human cardiovascular system antigen polypeptide SEQ ID No 1189.	56	27
1227	gi15029634	Homo sapiens	Similar to tetranectin (plasminogen-binding protein), clone MGC:13592 IMAGE:4042921, mRNA, complete cds.	725	100
1227	gi37409	Homo sapiens	H.sapiens mRNA for tetranectin.	725	100
1227	gi825722	Homo sapiens	H.sapiens tetranectin gene, exon 1.	725	100
1228	gi5790207	Taenia saginata	ATPase subunit 6	70	32
1228	gi3927873	Cyanidioschyzon merolae	NADH-ubiquinone oxidoreductase chain 3	44	19
1228	gi4378776	Pedinomonas minor	NADH dehydrogenase subunit 3	47	30
1229	AAE01790	Homo sapiens	HUMA- Human gene 21 encoded secreted protein HDPTW65, SEQ ID NO:111.	142	59
1229	AAE01838	Homo sapiens	HUMA- Human gene 21 encoded secreted protein HDPTW65, SEQ ID NO:159.	140	57
1229	ABB11479	Homo sapiens	HYSE- Human reverse transcriptase homologue, SEQ ID NO:1849.	92	55
1230	AAE04775	Homo sapiens	INCY- Human vesicle trafficking protein-18 (VETRP-18) protein.	327	100
1230	gi11120502	Homo sapiens	ERGL mRNA, complete cds.	327	100
1230	AAW88699	Homo sapiens	HUMA- Secreted protein encoded by gene 166 clone HCEQA68.	333	100
1231	AAG00381	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 4462.	266	91
1231	AAU19357	Homo sapiens	PHAA Human G protein-coupled receptor nGPCR-2290.	125	50
1231	AAO09238	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 23130.	109	75
1232	AAM06558	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 289.	301	98
1232	gi2648055	Xenopus laevis	corticotropin releasing factor receptor, type 2	48	29
1232	AAU21458	Homo sapiens	HUMA- Human novel foetal antigen, SEQ ID NO 1702.	45	36
1233	AAM06562	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 293.	383	100
1233	AAG72602	Homo sapiens	YEDA Human OR-like polypeptide query sequence, SEQ ID NO: 2283.	65	42
1233	gi7547265	Canis familiaris	5-Hydroxytryptamine 1D receptor subtype beta; 5-HT1D beta	67	39
1234	AAM92283	Homo sapiens	HUMA- Human digestive system antigen SEQ ID NO: 1632.	76	36
1234	AAO09955	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 23847.	70	35
1234	gi8778849	Arabidopsis	T7N9.1	69	42

Table 2A

162

SEQ ID	Hit ID	Species	Description	S score	% Identity
		thaliana			
1235	AAM63797	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 35902.	48	34
1235	AAM76610	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 36916.	48	34
1235	ABB12222	Homo sapiens	HYSE- Human secreted protein homologue, SEQ ID NO:2592.	52	42
1236	gi160822	Sarcocystis muris	31-kDa major surface antigen	55	37
1236	gi5305067	Mus musculus	immunoglobulin light chain variable region	61	34
1236	AAM60441	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 32546.	61	40
1237	gi4929790	Homo sapiens	angiopoietin-related protein 3 (ANGPTL3) mRNA, complete cds.	1489	98
1237	AAV05395	Homo sapiens	GETH Human TIE ligand NL6 protein sequence.	1489	98
1237	AAB12135	Homo sapiens	PROT- Hydrophobic domain protein from clone HP10622 isolated from Liver cells.	1489	98
1238	AAM06568	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 299.	142	57
1238	gi1334599	Magnaporthe grisea	ubiquinol:cytochrome c oxidoreductase	48	42
1238	gi13487283	Mycosphaerella fijiensis	cytochrome b	46	42
1239	gi15930235	Homo sapiens	clone IMAGE:3846772, mRNA, partial cds.	46	40
1239	gi1334235	Rattus rattus	MIP protein (261 AA; AA 3 - 263)	65	45
1239	gi1185419	Mus musculus	major intrinsic protein	65	45
1240	AAM60668	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 32773.	62	31
1240	AAM73340	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 33646.	62	31
1240	gi975678	Albinaria caerulea	start codon=CAT, termination codon=TAA	65	27
1241	AAG03454	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 7535.	40	26
1241	gi1245208	Danio rerio	Zg13	57	47
1241	AAV19486	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	33	41
1242	gi3157920	Arabidopsis thaliana	F12F1.6	46	37
1242	AAV48414	Homo sapiens	META- Human prostate cancer-associated protein 111.	36	46
1242	gi927722	Saccharomyces cerevisiae	Erd1p; YDR414C; CAI: 0.11	61	44
1243	gi9657469	Vibrio cholerae	soxR protein	39	46
1243	gi3493510	Mus musculus	Ymp	47	43
1243	gi2358254	Mus musculus	HNMP-1	47	43

Table 2A

163

SEQ ID	Hit ID	Species	Description	S score	% Identity
1244	AAO12129	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 26021.	64	90
1244	gi3874749	Caenorhabditis elegans	C34E7.3	50	56
1244	AAO12895	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 26787.	64	44
1245	AAV99386	Homo sapiens	GETH Human PRO1305 (UNQ671) amino acid sequence SEQ ID NO:153.	71	39
1245	AAO02040	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 15932.	63	33
1245	gi15073483	Sinorhizobium meliloti	PUTATIVE SENSORY TRANSDUCTION HISTIDINE KINASE TRANSMEMBRANE PROTEIN	77	39
1246	AAG75420	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:6184.	49	40
1246	gi4099021	Helicobacter pylori	amino acid permease	47	39
1246	gi2314328	Helicobacter pylori 26695	glutamine ABC transporter, permease protein (glnP)	47	39
1248	gi2959352	Brugia pahangi	cuticle collagen 2(F)	68	37
1248	AAO03627	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 17519.	59	40
1248	gi13959819	Myxococcus xanthus	USC6-1p	67	43
1249	gi8249629	Homo sapiens	partial IGHV gene for immunoglobulin heavy chain variable region, clone B31.	62	44
1249	gi6646882	Paragonimus westermani	NADH dehydrogenase subunit 1	63	40
1249	AAR39641	Homo sapiens	CIBA Transforming Growth Factor-beta1(44/45)beta2 hybrid.	44	37
1250	AAU04613	Homo sapiens	UNIW Gonadotropin analogue, beta subunit.	40	58
1250	gi3242155	Drosophila melanogaster	153C9.b	60	43
1250	AAM63639	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 35744.	59	42
1251	AAO11677	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 25569.	74	54
1251	gi7716782	Drosophila simulans	helicase pitchoune	77	44
1251	gi3342758	Drosophila melanogaster	helicase pitchoune	77	44
1252	gi482846	Torgos tracheliotus	cytochrome b	51	40
1252	gi22737	Hordeum vulgare	beta-hordothionin	42	38
1252	AAM79945	Homo sapiens	HYSE- Human protein SEQ ID NO 3591.	45	40
1253	gi424891	Human immunodeficiency virus type 1	envelope glycoprotein	37	33
1253	gi9654985	Vibrio cholerae	glutamate--cysteine ligase	62	28

Table 2A

164

SEQ ID	Hit ID	Species	Description	S score	% Identity
1254	gi3805960	Populus balsamifera subsp. trichocarpa	laccase	47	72
1254	gi15074016	Sinorhizobium meliloti	PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	57	35
1254	gi12652993	Homo sapiens	clone IMAGE:3357862, mRNA, partial cds.	56	47
1255	gi1655739	Peromyscus maniculatus	NADH dehydrogenase subunit 4	44	24
1255	gi16551105	Crotalus adamanteus	NADH dehydrogenase subunit 5	66	28
1255	gi16551107	Crotalus atrox	NADH dehydrogenase subunit 5	65	28
1256	AAO06799	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 20691.	43	37
1256	AAO00659	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 14551.	41	52
1256	AAB51937	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 9 SEQ ID NO:69.	37	60
1257	gi6449037	Mus musculus	platelet glycoprotein V	738	38
1257	gi2104856	Rattus norvegicus	platelet glycoprotein V	735	37
1257	gi2104845	Mus musculus	platelet glycoprotein V	722	37
1258	AAO11326	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 25218.	61	40
1258	gi7576909	Danio rerio	her7-protein	64	37
1258	AAG81428	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:374.	47	38
1259	gi18349	Daucus carota	glycine rich protein (AA 1 - 96)	65	45
1259	gi336034	Vesicular stomatitis virus	M-protein	70	26
1259	gi335876	Vesicular stomatitis virus	matrix (M) protein	70	26
1260	AAO09307	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 23199.	41	37
1260	gi15042581	Echinococcus granulosus	NADH dehydrogenase subunit 2	59	41
1261	gi10436454	Homo sapiens	cDNA FLJ14082 fis, clone HEMBB1002300.	983	99
1261	AAB95686	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18490.	983	99
1261	AAAY20668	Homo sapiens	UYRO- Human neurofilament-M wild type protein fragment 10.	44	50
1263	gi965014	Mus musculus	ADAM 4 protein precursor	1303	51
1263	gi1061159	Macaca fascicularis	testicular Metalloprotease-like, Disintegrin-like, Cysteine-rich protein IVa	1277	39
1263	gi1061161	Macaca fascicularis	testicular Metalloprotease-like, Disintegrin-like, Cysteine-rich protein IVb	1249	38
1264	AAM79049	Homo sapiens	HYSE- Human protein SEQ ID NO 1711.	1895	98
1264	AAM80033	Homo sapiens	HYSE- Human protein SEQ ID NO 3679.	1895	98
1264	AAM53458	Homo sapiens	MOLE- Human brain expressed	1074	100

Table 2A

165

SEQ ID	Hit ID	Species	Description	S score	% Identity
			single exon probe encoded protein SEQ ID NO: 25563.		
1265	AAB44605	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 10 SEQ ID NO:70.	93	70
1265	AAG71789	Homo sapiens	YEDA Human olfactory receptor polypeptide, SEQ ID NO: 1470.	42	63
1265	AAG72517	Homo sapiens	YEDA Human OR-like polypeptide query sequence, SEQ ID NO: 2198.	42	63
1266	gi14714741	Homo sapiens	clusterin (complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2, testosterone- repressed prostate message 2, apolipoprotein J), clone MGC:18080 IMAGE:4150452, mRNA, complete cds.	1629	99
1266	gi292843	Homo sapiens	Human TRPM-2 protein gene, exons 7,8,9 and complete cds.	1629	99
1266	gi30251	Homo sapiens	Human SP-40,40 mRNA for complement-associated protein SP- 40,40 alpha-1 and beta-1 chain.	1629	99
1267	gi11493504	Homo sapiens	PRO0309	1192	98
1267	gi412723	synthetic construct	synthetic antithrombin III	1192	98
1267	gi583741	synthetic construct	Antithrombin III	1192	98
1268	gi11493504	Homo sapiens	PRO0309	1439	98
1268	gi412723	synthetic construct	synthetic antithrombin III	1439	98
1268	gi583741	synthetic construct	Antithrombin III	1439	98
1269	gi203710	Rattus norvegicus	cytochrome c oxidase subunit VIc	250	65
1269	gi1200057	Homo sapiens	Human mRNA for cytochrome c oxidase subunit VIc.	229	61
1269	gi12652867	Homo sapiens	cytochrome c oxidase subunit VIc, clone MGC:1520 IMAGE:3350637, mRNA, complete cds.	229	61
1270	AAM96033	Homo sapiens	HUMA- Human reproductive system related antigen SEQ ID NO: 4691.	465	98
1270	AAU18881	Homo sapiens	HUMA- Novel prostate gland antigen, Seq ID No 180.	465	98
1270	gi9622236	Homo sapiens	cadherin-like protein VR20 mRNA, partial cds.	272	100
1271	gi552137	Drosophila melanogaster	tropomyosin isoform 9E	118	27
1271	gi158693	Drosophila melanogaster	tropomyosin isoform 9A	118	27
1271	gi158696	Drosophila melanogaster	tropomyosin isoform 9D	118	27
1272	gi6689249	Streptococcus dysgalactiae subsp. dysgalactiae	M-like protein	122	24

Table 2A  
166

SEQ ID	Hit ID	Species	Description	S score	% Identity
1272	gi6692690	Arabidopsis thaliana	F12K11.14	114	28
1272	gi4778	Saccharomyces cerevisiae	Uso1 protein	122	29
1273	gi13097573	Homo sapiens	Similar to thiosulfate sulfurtransferase (rhodanese), clone MGC:10492 IMAGE:3611253, mRNA, complete cds.	1469	94
1273	gi16876913	Homo sapiens	mercaptopyruvate sulfurtransferase, clone MGC:24539 IMAGE:4105509, mRNA, complete cds.	1469	94
1273	gi17511726	Homo sapiens	mercaptopyruvate sulfurtransferase, clone MGC:31798 IMAGE:4131927, mRNA, complete cds.	1469	94
1274	AAB85039	Homo sapiens	CURA- Human SER5 protein sequence.	767	48
1274	gi6137097	Homo sapiens	serine protease DESC1 (DESC1) mRNA, complete cds.	749	48
1274	AAAY99414	Homo sapiens	GETH Human PRO1461 (UNQ742) amino acid sequence SEQ ID NO:269.	749	48
1275	gi12584839	Homo sapiens	HT036-ISO (HT036-ISO) mRNA, complete cds.	997	94
1275	gi12584841	Homo sapiens	HT036 (HT036) mRNA, complete cds.	820	93
1275	gi17427028	Ralstonia solanacearum	CONSERVED HYPOTHETICAL PROTEIN	502	42
1276	gi310691	Simian virus 40	small T antigen	48	47
1276	gi8886685	Centris inermis	cytochrome b	53	40
1276	gi625084	Oncorhynchus tshawytscha	heat-shock protein 30	37	44
1277	gi7106820	Homo sapiens	HSPC215	261	100
1277	AAU16225	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1178.	261	100
1277	AAG81441	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:400.	261	100
1278	AAM25840	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1355.	208	88
1278	AAM74914	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 35220.	63	68
1278	AAM06639	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 370.	36	70
1279	gi7023943	Homo sapiens	GC36 mRNA, complete cds.	360	35
1279	gi2502077	Homo sapiens	digestive tract-specific calpain (nCL-4) mRNA, complete cds.	360	35
1279	gi2358262	Rattus norvegicus	calpain large subunit	351	35
1280	gi4153951	Homo sapiens	H.sapiens gene from PACs 295C6 and 313L4.	259	37
1280	AAAY32437	Homo sapiens	TEXA Absorptive hypercalciuria associated gene protein product.	259	37
1280	gi15383934	Homo sapiens	testicular soluble adenylyl cyclase mRNA, complete cds.	259	37

Table 2A  
167

SEQ ID	Hit ID	Species	Description	S score	% Identity
1281	AAM89651	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:17244.	66	31
1281	gi408591	Influenza A virus	nonstructural protein	62	28
1281	AAM82524	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:10117.	62	55
1282	gi4079820	Mus musculus	HERC2	67	40
1282	gi459017	Allomyces macrogynus	subunit 6 of the ATPase complex	71	44
1282	gi1236414	Allomyces macrogynus	H(+)-transporting ATPase, F0 subunit 6	71	44
1283	AAM63001	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 35106.	153	67
1283	AAM75812	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 36118.	153	67
1283	AAE10197	Homo sapiens	HYSE- Human bone marrow derived peptide, SEQ ID NO: 41.	60	36
1284	AAG81367	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:252.	816	98
1284	gi7582286	Homo sapiens	BM-007	530	98
1284	AAG02907	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6988.	408	95
1285	AAG81367	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:252.	906	98
1285	gi7582286	Homo sapiens	BM-007	538	99
1285	AAG02907	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6988.	416	96
1286	AAW49716	Homo sapiens	PROT- Protein polymer adhesive substrate PPAS1-C monomer.	54	31
1286	AAW49721	Homo sapiens	PROT- Protein polymer adhesive substrate PPAS1-D monomer.	54	31
1286	gi683735	Macaca fascicularis	endothelin 3	50	62
1287	gi5689766	Homo sapiens	mRNA for zinc finger 2 (ZNF2 gene).	2092	99
1287	gi14602980	Homo sapiens	clone MGC:16594 IMAGE:4110322, mRNA, complete cds.	1609	100
1287	gi13477207	Homo sapiens	clone MGC:12980 IMAGE:3350363, mRNA, complete cds.	1604	99
1288	AAB44228	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1673.	284	86
1288	AAM90208	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:17801.	43	50
1288	gi733438	Cepaea nemoralis	NADH dehydrogenase subunit 4L	62	31
1289	gi10764264	synthetic construct	mutated NS5A	67	30
1289	AAO02625	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 16517.	58	26

Table 2A  
168

SEQ ID	Hit ID	Species	Description	S score	% Identity
1289	gi10644188	Hepatitis C virus type 1a	polyprotein	67	30
1290	AAG03150	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 7231.	307	98
1290	AAW48931	Homo sapiens	CEDA- Schwannomin-binding protein C-terminal fragment.	286	100
1290	AAO04324	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 18216.	63	33
1291	AAB60098	Homo sapiens	INCY- Human transport protein TPPT-18.	1822	92
1291	gi1537070	Rattus norvegicus	nucleoporin p54	1767	92
1291	gi15214835	Homo sapiens	clone MGC:13407 IMAGE:3931652, mRNA, complete cds.	1822	92
1292	AAY94621	Homo sapiens	MILL- Epidermal growth factor-like variant in skin-2 amino acid sequence.	385	100
1292	AAE06697	Homo sapiens	HYSE- Human TGF alpha-like protein.	385	100
1292	AAE06698	Homo sapiens	HYSE- Human TGF alpha-like splice variant protein.	385	100
1293	AAW78245	Homo sapiens	HUMA- Fragment of human secreted protein encoded by gene 19.	1018	98
1293	ABB11835	Homo sapiens	HYSE- Human secreted protein homologue, SEQ ID NO:2205.	1018	98
1293	AAM79352	Homo sapiens	HYSE- Human protein SEQ ID NO 2998.	1018	98
1294	AAM99920	Homo sapiens	HUMA- Human polypeptide SEQ ID NO 36.	667	97
1294	gi16552010	Homo sapiens	cDNA FLJ32009 fis, clone NT2RP7009498, weakly similar to FIBULIN-1, ISOFORM A PRECURSOR.	667	97
1294	AAM99933	Homo sapiens	HUMA- Human polypeptide SEQ ID NO 49.	627	93
1295	gi2598167	Homo sapiens	zinc finger protein (HZF6) mRNA, 5' UTR and partial cds.	2772	99
1295	gi5640019	Mus musculus	zinc finger protein ZFP235	1565	68
1295	gi1184371	Mus musculus	zinc finger protein; Method: conceptual translation supplied by author	1278	55
1296	gi15679947	Homo sapiens	endothelial zinc finger protein induced by tumor necrosis factor alpha, clone MGC:11153 IMAGE:3840512, mRNA, complete cds.	2734	100
1296	gi9502202	Homo sapiens	endothelial zinc finger protein induced by tumor necrosis factor alpha (EZFIT) mRNA, complete cds.	2734	100
1296	gi10437767	Homo sapiens	cDNA: FLJ21628 fis, clone COL08076.	1713	77
1297	AAM56742	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 28847.	99	55



Table 2A  
169

SEQ ID	Hit ID	Species	Description	S score	% Identity
1297	AAO09197	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 23089.	66	45
1297	gi12543402	Corynebacterium glutamicum	FRXA00675	79	26
1298	AAM79176	Homo sapiens	HYSE- Human protein SEQ ID NO 1838.	601	100
1298	ABB11626	Homo sapiens	HYSE- Human Fas-associated phosphatase homologue, SEQ ID NO:1996.	559	94
1298	AAM80160	Homo sapiens	HYSE- Human protein SEQ ID NO 3806.	559	94
1299	gi12698338	Homo sapiens	matrix metalloproteinase-28 precursor, mRNA, complete cds.	2424	96
1299	gi12698852	Homo sapiens	matrix metalloprotease MMP25 mRNA, complete cds.	2424	96
1299	AAU12243	Homo sapiens	GETH Human PRO4339 polypeptide sequence.	2424	96
1300	gi14210477	Homo sapiens	interleukin 18 precursor, mRNA, complete cds.	138	92
1300	AAW31757	Homo sapiens	INCY- Interferon gamma inducing factor-2 (IGIF-2) R140I variant.	138	92
1300	gi10799833	Ovis aries	interleukin-18 (IGIF)	122	78
1301	AAE05302	Homo sapiens	MILL- Human TANGO 457 protein.	623	97
1301	AAE05303	Homo sapiens	MILL- Human mature TANGO 457 protein.	611	100
1301	AAE05305	Homo sapiens	MILL- Human TANGO 457 protein cytoplasmic domain.	605	100
1302	AAM55396	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 27501.	64	38
1302	AAM57742	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 29847.	64	38
1302	AAM67792	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 28098.	64	38
1303	AAM88370	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:15963.	38	35
1303	gi7330034	Macaca mulatta rhadinovirus 26-95	helicase-primase	56	30
1303	gi4494949	Macaca mulatta rhadinovirus 17577	helicase/primase	56	30
1304	gi190870	Homo sapiens	Human retinoic acid receptor gamma 2 mRNA, 5' end.	274	100
1304	gi297146	Homo sapiens	H.sapiens gene for retinoic acid receptor gamma-2.	274	100
1304	gi200660	Mus musculus	retinoic acid receptor gamma 2	252	92
1305	AAM39737	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2882.	992	99
1305	AAM39736	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2881.	875	100
1305	AAM41522	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 6453.	875	100

Table 2A

170

SEQ ID	Hit ID	Species	Description	S score	% Identity
1306	ABB17891	Homo sapiens	HUMA- Human nervous system related polypeptide SEQ ID NO 6548.	54	38
1306	AAM88996	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:16589.	57	40
1306	AAM65093	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 37198.	55	50
1307	AAY19551	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	133	42
1307	AAY75972	Homo sapiens	GENE- Human skin cell protein, SEQ ID 150.	133	42
1307	AAB55911	Homo sapiens	GENE- Skin cell protein, SEQ ID NO: 150.	133	42
1308	AAU27671	Homo sapiens	ZYMO Human protein AFP355471.	486	100
1308	AAO12566	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 26458.	61	28
1308	gi167933	Dictyostelium discoideum	thioredoxin	67	34
1309	AAR15222	Homo sapiens	TEXA Chronic myelogenous leukaemia-derived myeloid-related protein.	620	100
1309	gi181527	Homo sapiens	Human neutrophil peptide (defensin) 1 mRNA, complete cds.	493	100
1309	gi181529	Homo sapiens	Human defensin 1 protein mRNA, complete cds.	493	100
1310	gi2911559	Human papillomavirus type 77	E6 protein	66	27
1310	gi9800324	rat cytomegalovirus Maastricht	pr109	62	25
1310	gi397007	Human papillomavirus type 3	envelope protein	60	31
1311	AAU19632	Homo sapiens	HUMA- Human novel extracellular matrix protein, Seq ID No 282.	205	40
1311	gi3127926	Homo sapiens	H.sapiens RNA for type VI collagen alpha3 chain.	186	42
1311	gi57960	Mus musculus	collagen alpha 3 chain type VI	176	40
1312	gi16508176	Homo sapiens	small GTP-binding tumor suppressor 1 mRNA, complete cds.	1012	100
1312	gi16555334	Homo sapiens	Rig protein mRNA, complete cds.	1012	100
1312	gi16508174	Mus musculus	small GTP-binding tumor suppressor 1	963	93
1313	AAG73984	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:4748.	55	40
1313	gi3041771	Homo sapiens	mRNA for perilipin, complete cds.	83	32
1313	AAY22157	Homo sapiens	ABBO Human BS135 protein sequence.	83	32
1314	AAM71801	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 32107.	1872	100
1314	gi16549907	Homo sapiens	cDNA FLJ30663 fis, clone	1203	58

Table 2A

171

SEQ ID	Hit ID	Species	Description	S score	% Identity
			FCBBF1000598, moderately similar to ZINC FINGER PROTEIN 84.		
1314	AAM78565	Homo sapiens	HYSE- Human protein SEQ ID NO 1227.	1151	55
1315	gi32472	Homo sapiens	H.sapiens mRNA for high-sulphur keratin.	785	76
1315	gi3228239	Homo sapiens	UHS KerA gene.	774	76
1315	gi34079	Homo sapiens	Human gene for ultra high-sulphur keratin protein.	774	76
1316	gi12655446	Homo sapiens	mRNA for keratin associated protein 4.4 (KRTAP4.4 gene).	755	80
1316	gi12655460	Homo sapiens	mRNA for keratin associated protein 4.12 (KRTAP4.12 gene).	726	75
1316	gi13278825	Homo sapiens	Similar to RIKEN cDNA 1110054P19 gene, clone MGC:2782 IMAGE:2959821, mRNA, complete cds.	726	75
1317	gi12655462	Homo sapiens	mRNA for keratin associated protein 4.14 (KRTAP4.14 gene).	1102	88
1317	gi12655452	Homo sapiens	mRNA for keratin associated protein 4.7 (KRTAP4.7 gene).	1081	84
1317	gi12655460	Homo sapiens	mRNA for keratin associated protein 4.12 (KRTAP4.12 gene).	997	79
1318	AAM79404	Homo sapiens	HYSE- Human protein SEQ ID NO 3050.	850	74
1318	AAM39466	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2611.	844	77
1318	AAM41252	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 6183.	836	77
1319	gi16552172	Homo sapiens	cDNA FLJ32133 fis, clone PEBLM2000308, moderately similar to ZINC FINGER PROTEIN 135.	800	65
1319	gi6467200	Homo sapiens	GIOT-1 mRNA for gonadotropin inducible transcription repressor-1, partial cds.	775	60
1319	gi498721	Homo sapiens	H.sapiens HZF10 mRNA for zinc finger protein.	770	63
1320	gi3036963	Ciona savignyi	CsCDC42	163	60
1320	gi15072535	Schizophyllum commune	small GTPase CDC42	162	60
1320	gi520533	Drosophila melanogaster	Dcdc42	161	60
1321	AAE02058	Homo sapiens	HUMA- Human four disulfide core domain (FDCD)-containing protein.	517	43
1321	gi12655452	Homo sapiens	mRNA for keratin associated protein 4.7 (KRTAP4.7 gene).	509	44
1321	gi200964	Mus musculus	serine 2 ultra high sulfur protein	494	42
1322	ABB12490	Homo sapiens	HYSE- Human bone marrow expressed protein SEQ ID NO: 329.	169	72
1322	gi14647047	Puntius titteya	ATP synthase 8	56	37
1322	gi14646929	Barbus cyclolepis	ATP synthase 8	54	44

Table 2A  
172

SEQ ID	Hit ID	Species	Description	S score	% Identity
1323	gi5921470	Homo sapiens	mRNA for G8 protein (G8 gene, located in the class III region of the major histocompatibility complex).	405	89
1323	gi5921473	Homo sapiens	mRNA for G8 protein (G8 gene, located in the class III region of the major histocompatibility complex), alternative splice variant lacking exon 2.	381	92
1323	AAM39144	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2289.	381	92
1324	AAM87150	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:14743.	74	38
1324	AAB59115	Homo sapiens	HUMA- Breast and ovarian cancer associated antigen protein sequence SEQ ID 823.	62	36
1324	gi15158712	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_L_725p	46	52
1325	AAV48404	Homo sapiens	META- Human prostate cancer-associated protein 101.	55	50
1325	AAM59935	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 32040.	55	38
1325	AAM72530	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 32836.	55	38
1326	gi466912	Mycobacterium leprae	tp2; B1549_C2_206	60	45
1326	gi1220377	Avian infectious bronchitis virus	nucleocapsid protein	64	56
1326	gi13177409	Ectocarpus siliculosus virus	EsV-1-135	65	34
1327	gi200964	Mus musculus	serine 2 ultra high sulfur protein	207	32
1327	gi200962	Mus musculus	serine 1 ultra high sulfur protein	202	32
1327	gi32472	Homo sapiens	H.sapiens mRNA for high-sulphur keratin.	196	32
1328	AAR23732	Homo sapiens	MINU Gene 519 cDNA derived peptide.	316	68
1328	gi35065	Homo sapiens	Human NKG5 mRNA, expressed in natural killer cells and T-cells.	314	66
1328	AAW59874	Homo sapiens	HUMA- Amino acid sequence of the cDNA clone CAT-1 (HTXET53).	314	66
1329	gi200964	Mus musculus	serine 2 ultra high sulfur protein	359	49
1329	AAM39466	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 2611.	341	51
1329	AAM78420	Homo sapiens	HYSE- Human protein SEQ ID NO 1082.	337	53
1330	gi13937769	Homo sapiens	Similar to RIKEN cDNA 1200013F24 gene, clone MGC:12197 IMAGE:3997840, mRNA, complete cds.	1256	100
1330	gi7582294	Homo sapiens	BM-011	781	98
1330	AAM79664	Homo sapiens	HYSE- Human protein SEQ ID NO 3310.	255	31
1331	gi14718451	Homo sapiens	sialic acid-binding lectin 11	796	71

Table 2A

173

SEQ ID	Hit ID	Species	Description	S score	% Identity
			(SIGLEC11) mRNA, complete cds.		
1331	AAY41724	Homo sapiens	GETH Human PRO940 protein sequence.	535	50
1331	AAB44280	Homo sapiens	GETH Human PRO940 (UNQ477) protein sequence SEQ ID NO:259.	535	50
1332	AAT90731_aal	Homo sapiens	FARB Human placental bikunin cDNA.	849	98
1332	gi12804515	Homo sapiens	serine protease inhibitor, Kunitz type, 2, clone MGC:2021 IMAGE:2959462, mRNA, complete cds.	848	98
1332	gi2065529	Homo sapiens	placental bikunin mRNA, complete cds.	848	98
1333	gi14042550	Homo sapiens	cDNA FLJ14779 fis, clone NT2RP4000398, moderately similar to ZINC FINGER PROTEIN 140.	2165	98
1333	AAB93164	Homo sapiens	HELI- Human protein sequence SEQ ID NO:12091.	2165	98
1333	AAM93693	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3604.	2159	100
1334	gi12804907	Homo sapiens	Similar to metaxin 1, clone MGC:2518 IMAGE:3546178, mRNA, complete cds.	1512	100
1334	gi1326108	Homo sapiens	Human metaxin (MTX) gene, complete cds.	1098	100
1334	gi2564913	Homo sapiens	clk2 kinase (CLK2), propin1, cotel1, glucocerebrosidase (GBA), and metaxin genes, complete cds; metaxin pseudogene and glucocerebrosidase pseudogene; and thrombospondin3 (THBS3) gene, partial cds.	1098	100
1335	AAW85614	Homo sapiens	GEMY Secreted protein clone fr473_2.	381	83
1335	AAY94865	Homo sapiens	PROT- Human protein clone HP10540.	381	83
1335	AAY36022	Homo sapiens	GEST Extended human secreted protein sequence, SEQ ID NO. 407.	365	80
1336	AAB18447	Homo sapiens	MILL- Amino acid sequence of human TANGO 216 polypeptide.	2257	99
1336	AAB18455	Homo sapiens	MILL- A human TANGO 216 polypeptide clone.	2257	99
1336	AAU19662	Homo sapiens	HUMA- Human novel extracellular matrix protein, Seq ID No 312.	1876	96
1337	AAB18447	Homo sapiens	MILL- Amino acid sequence of human TANGO 216 polypeptide.	2257	99
1337	AAB18455	Homo sapiens	MILL- A human TANGO 216 polypeptide clone.	2257	99
1337	AAU19662	Homo sapiens	HUMA- Human novel extracellular matrix protein, Seq ID No 312.	1876	96
1338	AAY86303	Homo sapiens	HUMA- Human secreted protein HOGCK20, SEQ ID NO:218.	2133	94

Table 2A

174

SEQ ID	Hit ID	Species	Description	S score	% Identity
1338	gi14456615	Homo sapiens	PIG-T mRNA for phosphatidyl inositol glycan class T, complete cds.	2120	96
1338	gi15929132	Homo sapiens	clone MGC:8909 IMAGE:3921680, mRNA, complete cds.	2120	96
1339	gi12836893	Gallus gallus	IPR328-like protein	160	29
1339	gi3093433	Homo sapiens	Chromosome 16 BAC clone CIT987SK-625P11, complete sequence.	153	29
1339	gi4558766	Homo sapiens	neuronal voltage gated calcium channel gamma-3 subunit mRNA, complete cds.	153	29
1340	gi12836893	Gallus gallus	IPR328-like protein	158	29
1340	gi3093433	Homo sapiens	Chromosome 16 BAC clone CIT987SK-625P11, complete sequence.	151	29
1340	gi4558766	Homo sapiens	neuronal voltage gated calcium channel gamma-3 subunit mRNA, complete cds.	151	29
1341	AAW85737	Homo sapiens	SAGA Polypeptide with transmembrane domain.	692	100
1341	ABB11882	Homo sapiens	HYSE- Human transmembrane protein homologue, SEQ ID NO:2252.	692	100
1341	AAG89353	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 473.	692	100
1342	ABB12032	Homo sapiens	HYSE- Human SIGP 2328134 homologue, SEQ ID NO:2402.	866	97
1342	AAY21851	Homo sapiens	INCY- Human signal peptide-containing protein (SIGP) (clone ID 2328134).	866	97
1342	gi4101574	Homo sapiens	54TMp (54tm) mRNA, complete cds.	860	96
1343	gi3002925	Homo sapiens	T cell receptor beta chain (TCRBV13S1-TCRBJ2S1) mRNA, complete cds.	1658	100
1343	gi2982508	Homo sapiens	mRNA for TCR beta chain, specific for Mage 3/HLA-A2.	1527	94
1343	gi3002933	Homo sapiens	T cell receptor beta chain (TCRBV3S1-TCRBJ2S3) mRNA, complete cds.	1251	76
1344	gi14973269	Streptococcus pneumoniae TIGR4	cell wall surface anchor family protein	481	19
1344	gi15991793	Streptococcus gordonii	platelet binding protein GspB	303	17
1344	gi8885520	Streptococcus gordonii	streptococcal hemagglutinin	293	16
1345	AAY07751	Homo sapiens	HUMA- Human secreted protein fragment encoded from gene 8.	293	100
1345	gi1142588	Trypanosoma brucei	CR3	85	42
1345	gi3037018	Bodo saltans	NADH dehydrogenase subunit 5	80	35
1346	AAG78000	Homo sapiens	BIOW- Human actin 14.	663	100
1346	ABB17913	Homo sapiens	HUMA- Human nervous system related polypeptide SEQ ID NO	644	98

Table 2A

175

SEQ ID	Hit ID	Species	Description	S score	% Identity
			6570.		
1346	AAAY94954	Homo sapiens	GEMY Human secreted protein clone iw66_1 protein sequence SEQ ID NO:114.	596	68
1347	gi9837433	Homo sapiens	sialic acid binding immunoglobulin-like lectin 8 long splice variant (Siglec8) gene, complete cds.	2206	88
1347	gi6746556	Homo sapiens	sialic acid-binding immunoglobulin-like lectin-8 (SIGLEC8) mRNA, complete cds.	2031	93
1347	gi6980022	Homo sapiens	siglec SAF2 (SAF2) mRNA, complete cds.	2031	93
1348	gi15451469	Homo sapiens	siglec-like protein (SLG2) gene and alternatively spliced variants, complete cds.	2689	99
1348	gi15217166	Homo sapiens	sialic acid-binding Ig-like lectin 10 (SIGLEC10) mRNA, complete cds.	2682	99
1348	gi14164613	Homo sapiens	sialic acid binding immunoglobulin-like lectin 10 (SIGLEC10) mRNA, complete cds.	2356	98
1349	AAB60112	Homo sapiens	INCY- Human transport protein TPPT-32.	775	100
1349	gi9663117	Homo sapiens	mRNA for organic cation transporter.	382	48
1349	AAB47000	Homo sapiens	BOEN/ Human BOCT protein.	382	48
1350	AAAY76219	Homo sapiens	HUMA- Human secreted protein encoded by gene 96.	336	94
1350	gi2906006	Homo sapiens	WASP interacting protein (WIP) mRNA, partial cds.	134	30
1350	gi22269	Zea mays	cell wall protein (108 AA)	105	34
1351	AAB08767	Homo sapiens	INCY- A human leukocyte and blood related protein (LBAP).	92	37
1351	gi576631	Torpedo marmorata	14 kDa transmembrane protein	87	32
1351	AAM78542	Homo sapiens	HYSE- Human protein SEQ ID NO 1204.	72	31
1352	gi5817194	Homo sapiens	mRNA; cDNA DKFZp434F011 (from clone DKFZp434F011); partial cds.	221	95
1352	gi7576452	Homo sapiens	hBOIT mRNA for potent brain type organic ion transporter, complete cds.	160	36
1352	AAAY06116	Homo sapiens	MILL- Human organic cation transporter OCT-3.	154	35
1353	gi16552104	Homo sapiens	cDNA FLJ32082 fis, clone OCBBF2000231, weakly similar to PHOSPHOLIPASE A2 INHIBITOR SUBUNIT B PRECURSOR.	2566	99
1353	AAAY66713	Homo sapiens	GETH Membrane-bound protein PRO1309.	2566	99
1353	AAU12396	Homo sapiens	GETH Human PRO1309 polypeptide sequence.	2566	99
1354	gi15559274	Homo sapiens	clone MGC:20205	502	40

Table 2A  
176

SEQ ID	Hit ID	Species	Description	S score	% Identity
			IMAGE:3627858, mRNA, complete cds.		
1354	AAB15549	Homo sapiens	INCY- Human immune system molecule from Incyte clone 2774913.	502	40
1354	AAB19729	Homo sapiens	CURA- Human SECX Clone 4339264-2 encoded protein.	502	40
1355	AAAY99399	Homo sapiens	GETH Human PRO1268 (UNQ638) amino acid sequence SEQ ID NO:214.	603	100
1355	AAAY78808	Homo sapiens	PROT- Hydrophobic domain containing protein clone HP10537 protein sequence.	603	100
1355	AAB87570	Homo sapiens	GETH Human PRO1268.	603	100
1356	AAM78418	Homo sapiens	HYSE- Human protein SEQ ID NO 1080.	1902	97
1356	ABB11897	Homo sapiens	HYSE- Human F22162_1 homologue, SEQ ID NO:2267.	1827	93
1356	AAM79402	Homo sapiens	HYSE- Human protein SEQ ID NO 3048.	1820	93
1357	gi397607	Homo sapiens	H.sapiens encoding CLA-1 mRNA.	2331	99
1357	AAAY49573	Homo sapiens	WHED Human CLA-1 protein sequence.	2331	99
1357	AAW97900	Homo sapiens	MILL- Human SR-BI class B scavenger.	2318	98
1358	gi854065	Human herpesvirus 6	U88	348	31
1358	gi10434098	Homo sapiens	cDNA FLJ12547 fis, clone NT2RM4000634.	273	32
1358	AAB95124	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17122.	273	32
1359	AAE05302	Homo sapiens	MILL- Human TANGO 457 protein.	1521	96
1359	AAE05303	Homo sapiens	MILL- Human mature TANGO 457 protein.	1397	96
1359	AAE05305	Homo sapiens	MILL- Human TANGO 457 protein cytoplasmic domain.	1260	100
1360	gi10129690	Homo sapiens	mRNA for mucolipidin (ML4 gene).	804	53
1360	gi10438844	Homo sapiens	cDNA: FLJ22449 fis, clone HRC09609.	804	53
1360	gi13477347	Homo sapiens	mucolipin 1, clone MGC:3287 IMAGE:3507836, mRNA, complete cds.	804	53
1361	AAE04122	Homo sapiens	HUMA- Human gene 23 encoded secreted protein HE8OK73, SEQ ID NO:108.	214	61
1361	AAE04169	Homo sapiens	HUMA- Human gene 23 encoded secreted protein HE8OK73, SEQ ID NO:158.	207	60
1361	AAG00392	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 4473.	117	43
1362	AAAY27853	Homo sapiens	HUMA- Human secreted protein encoded by gene No. 101.	274	94
1362	gi904289	Phaseolus vulgaris	fungine endopolygalacturonase inhibitor	66	41



Table 2A

177

SEQ ID	Hit ID	Species	Description	S score	% Identity
1362	AAB27594	Homo sapiens	HUMA- Human secreted protein SEQ ID NO: 95.	60	58
1363	gi17016967	Homo sapiens	NUANCE (NUA) mRNA, complete cds; alternatively spliced.	3404	71
1363	gi17861384	Homo sapiens	nesprin-2 gamma mRNA, complete cds.	3404	71
1363	gi5262574	Homo sapiens	mRNA; cDNA DKFZp434G173 (from clone DKFZp434G173); complete cds.	3404	71
1364	AAB95854	Homo sapiens	HELI- Human protein sequence SEQ ID NO:18912.	72	37
1364	gi9621943	Pelargonium senecioides	NADH dehydrogenase	58	42
1364	gi9621945	Pelargonium trifidum	NADH dehydrogenase	58	42
1365	AAW29654	Homo sapiens	GEMY Human secreted protein DM406_1.	140	48
1365	gi10187870	Rhodococcus sp.	ohpA transport	75	26
1365	gi15559671	Homo sapiens	clone MGC:20633 IMAGE:4761663, mRNA, complete cds.	72	33
1366	gi10566471	Mus musculus	Gliacolin	850	68
1366	gi14278927	Mus musculus	gliacolin	850	68
1366	gi3747099	Mus musculus	Clq-related factor	724	66
1367	gi2745756	Aotus trivirgatus	ribonuclease k6 precursor	431	91
1367	gi5730384	Eulemur fulvus collaris	ribonuclease k6 precursor	305	86
1367	gi5730382	Nycticebus coucang	ribonuclease k6 precursor	279	80
1368	AAE09651	Homo sapiens	HUMA- Human gene 13 encoded lipid metabolism protein HTJN173, SEQ ID NO:45.	484	98
1368	AAG64355	Homo sapiens	UYFU- Human lambda crystallin.	400	97
1368	AAAY92506	Homo sapiens	INCY- Human OXRE-3 with identity to lambda crystallin.	381	98
1369	AAM25241	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:756.	484	95
1369	AAE09651	Homo sapiens	HUMA- Human gene 13 encoded lipid metabolism protein HTJN173, SEQ ID NO:45.	352	100
1369	AAG64355	Homo sapiens	UYFU- Human lambda crystallin.	268	98
1370	AAM79626	Homo sapiens	HYSE- Human protein SEQ ID NO 3272.	214	100
1370	AAM79368	Homo sapiens	HYSE- Human protein SEQ ID NO 3014.	70	44
1370	AAO02702	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 16594.	65	36
1371	gi6653659	Oryctolagus cuniculus	chloride channel CLC-6	3919	96
1371	ABB11826	Homo sapiens	HYSE- Human Cl channel homologue, SEQ ID NO:2196.	3865	96
1371	gi1770376	Homo sapiens	H.sapiens mRNA for chloride channel, CLC-6a.	1620	100
1372	AAG71967	Homo sapiens	YEDA Human olfactory receptor polypeptide, SEQ ID NO: 1648.	725	97
1372	AAG71962	Homo sapiens	YEDA Human olfactory receptor	714	95

Table 2A

178

SEQ ID	Hit ID	Species	Description	S score	% Identity
			polypeptide, SEQ ID NO: 1643.		
1372	gi12007416	Mus musculus	m51 olfactory receptor	553	72
1373	AAM63071	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 35176.	278	100
1373	AAM75882	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 36188.	278	100
1373	AAM67333	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 27639.	234	100
1374	AAH26531_aal	Homo sapiens	REGC Human proton/oligonucleotide transporter hPHT1 cDNA.	920	100
1374	AAB82821	Homo sapiens	REGC Human proton/oligonucleotide transporter hPHT1 polypeptide.	815	100
1374	gi2208839	Rattus norvegicus	peptide/histidine transporter	721	82
1375	gi190418	Homo sapiens	Human cathepsin L gene, complete cds.	1597	87
1375	gi29715	Homo sapiens	Human mRNA for pro-cathepsin L (major excreted protein MEP).	1597	87
1375	AAW47031	Homo sapiens	USSH Human procathepsin L.	1597	87
1376	gi16566341	Homo sapiens	G protein-coupled receptor (GPR101) gene, complete cds.	2554	98
1376	AAB86428	Homo sapiens	BOEN/ Human brain SERALPHA protein.	2554	98
1376	AAU04369	Homo sapiens	AREN- Human G-protein coupled receptor, hRUP15.	2554	98
1377	AAY53605	Homo sapiens	METR- Peptide of human KChAP that binds to KValpha and Kvbeta subunits.	79	35
1377	AAY53608	Homo sapiens	METR- KChAP domain that binds to KValpha and Kvbeta subunits.	78	35
1377	gi6102853	Homo sapiens	mRNA; cDNA DKFZp727A051 (from clone DKFZp727A051); partial cds.	82	38
1378	AAB61616	Homo sapiens	PROT- Human protein HP10678.	2416	94
1378	AAG68126	Homo sapiens	FARB Human 7TM-GPCR protein sequence SEQ ID NO:6.	966	40
1378	AAE12023	Homo sapiens	INCY- Human G-protein coupled receptor, GCREC-2.	951	40
1379	AAY30735	Homo sapiens	HUMA- Amino acid sequence of a human secreted protein.	280	100
1379	gi333947	Human respiratory syncytial virus	membrane glycoprotein	48	39
1379	gi222567	Human respiratory syncytial virus	SH protein	48	39
1380	gi2459682	Homo sapiens	MAGE-B2 (MAGE-B2), MAGE-B3 (MAGE-B3), MAGE-B4 (MAGE-B4), and MAGE-B1 (MAGE-B1) genes, complete cds.	884	55
1380	gi3687199	Homo sapiens	Xp22 bins 169-171 BAC GSHB-383H3 (Genome Systems Human	884	55

Table 2A  
179

SEQ ID	Hit ID	Species	Description	S score	% Identity
			BAC Library) complete sequence.		
1380	gi4033512	Homo sapiens	DAM10 exon 3, partial sequence; and DAM10=DSS-AHC critical interval MAGE superfamily protein (DAM10) gene, complete cds.	867	54
1381	gi10198115	Homo sapiens	2P domain potassium channel TREK2 (KCNK10) mRNA, complete cds.	2697	100
1381	gi8452900	Rattus norvegicus	potassium channel TREK-2	2556	95
1381	gi4584799	Mus musculus	TREK-1 K+ channel subunit	1244	65
1382	ABB11297	Homo sapiens	HYSE- Human Coxsackie adenovirus receptor homologue, SEQ ID NO:1667.	699	97
1382	gi14279421	Danio rerio	coxsackievirus and adenovirus receptor-like protein	312	40
1382	gi6013133	Rattus norvegicus	coxsackie-adenovirus-receptor homolog	306	39
1383	gi17016394	Homo sapiens	cervical cancer 1 proto-oncogene-binding protein KG19 (KG19) mRNA, complete cds.	753	100
1383	AAM58441	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 30546.	99	28
1383	AAB86126	Homo sapiens	HUMA- Human MIT-like protein fragment encoded by cDNA clone HMSMI80.	99	28
1384	AAM06866	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 1074.	1133	95
1384	gi15099951	Mus musculus	diacylglycerol acyltransferase 2	959	51
1384	gi15099953	Homo sapiens	diacylglycerol acyltransferase 2 mRNA, complete cds.	951	50
1386	gi338506	Homo sapiens	Human salivary statherin gene, exons 2-6.	254	83
1386	gi338508	Homo sapiens	Human statherin mRNA, complete cds.	254	83
1386	AAY94527	Homo sapiens	INCY- Human statherin protein.	254	83
1387	gi10435784	Homo sapiens	cDNA FLJ13693 fis, clone PLACE2000111.	1011	100
1387	AAB94721	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15739.	1011	100
1387	gi13592427	Caenorhabditis elegans	similar to glycoproteins	124	34
1388	gi12654579	Homo sapiens	peptidylprolyl isomerase B (cyclophilin B), clone MGC:2224 IMAGE:2966791, mRNA, complete cds.	918	98
1388	gi14250758	Homo sapiens	peptidylprolyl isomerase B (cyclophilin B), clone MGC:14109 IMAGE:3502055, mRNA, complete cds.	918	98
1388	gi337999	Homo sapiens	Human secreted cyclophilin-like protein (SCYLP) mRNA, complete cds.	918	98
1389	AAE07112	Homo sapiens	HUMA- Human gene 6 encoded secreted protein fragment, SEQ ID	2473	99

Table 2A  
180

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NO:129.		
1389	gi9368530	Homo sapiens	mRNA full length insert cDNA clone EUROIMAGE 363668.	2457	99
1389	gi12053163	Homo sapiens	mRNA; cDNA DKFZp434D0727 (from clone DKFZp434D0727); complete cds.	2378	99
1390	gi16589056	Homo sapiens	type II gonadotropin-releasing hormone receptor gene, partial cds.	1021	99
1390	gi17048804	Homo sapiens	The CDS shown includes some apparent amino acids (from Gly 10 onwards) which would be deleted in a short intron	1008	98
1390	gi14029600	Cercopithecus aethiops	GnRH receptor II	944	92
1391	gi16359249	Mus musculus	RIKEN cDNA 1300010M03 gene	2226	91
1391	AAM93450	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3100.	575	37
1391	gi10438431	Homo sapiens	cDNA: FLJ22155 fis, clone HRC00205.	596	34
1392	AAE04896	Homo sapiens	INCY- Human transporter and ion channel-9 (TRICH-9) protein.	825	100
1392	gi12003980	Homo sapiens	spinster-like protein mRNA, complete cds.	695	52
1392	gi14249892	Homo sapiens	spinster-like protein, clone MGC:15767 IMAGE:3501826, mRNA, complete cds.	695	52
1393	AAG71515	Homo sapiens	YEDA Human olfactory receptor polypeptide, SEQ ID NO: 1196.	1051	94
1393	AAG72603	Homo sapiens	YEDA Human OR-like polypeptide query sequence, SEQ ID NO: 2284.	1051	94
1393	AAU24762	Homo sapiens	SENO- Human olfactory receptor AOLFR130B.	482	47
1394	AAB08894	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 4 SEQ ID NO:51.	165	59
1394	gi15626257	Buffalopox virus	p8 protein homologue	69	40
1394	AAO06451	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 20343.	69	35
1395	gi2792525	Equus caballus	connexin 43	64	35
1395	gi15148992	Human immunodeficiency virus type 1	vpu protein	64	30
1395	gi5738572	Human immunodeficiency virus type 1	VPU protein	60	33
1396	AAM83617	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:11210.	61	41
1396	gi4467773	Helicobacter pylori	cytotoxin associated protein A	60	34
1396	gi7248699	Helicobacter pylori	cytotoxin associated protein CagA	60	34
1397	gi11862939	Mus musculus	DDM36	5233	88
1397	gi11862941	Mus musculus	DDM36E	5224	88
1397	gi7650186	Mus musculus	neighbor of Punc e11 protein	5196	87

Table 2A

181

SEQ ID	Hit ID	Species	Description	S score	% Identity
1398	gi5596705	Homo sapiens	Novel human mRNA similar to C. elegans gene WP:CE18674, TR:Q19985.	1305	71
1398	gi15292481	Drosophila melanogaster	SD03655p	1174	59
1398	AAB88372	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0108.	886	71
1399	gi1335598	Simian sarcoma virus	coding sequence of p15E	64	33
1399	gi14039584	Casuaris casuaris	ATPase 8	55	50
1399	gi17427567	Ralstonia solanacearum	HYPOTHETICAL TRANSMEMBRANE PROTEIN	72	42
1400	AAM88491	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:16084.	38	41
1400	AAO09674	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 23566.	38	44
1400	gi3845106	Plasmodium falciparum	metal binding protein (DHHC domain)	55	36
1401	gi10434098	Homo sapiens	cDNA FLJ12547 fis, clone NT2RM4000634.	149	34
1401	AAB95124	Homo sapiens	HELI- Human protein sequence SEQ ID NO:17122.	149	34
1401	AAO09309	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 23201.	139	26
1402	AAE01249	Homo sapiens	HUMA- Human gene 18 encoded secreted protein HFIIN69, SEQ ID NO:111.	222	100
1402	AAE01299	Homo sapiens	HUMA- Human gene 18 encoded secreted protein HFIIN69, SEQ ID NO:162.	222	100
1402	AAE01332	Homo sapiens	HUMA- Human gene 18 encoded secreted protein fragment, SEQ ID NO:197.	222	100
1403	AAM06589	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 320.	237	100
1403	gi10732779	Mus musculus	APRIL	56	43
1403	AAM87662	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:15255.	38	58
1404	AAO09486	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 23378.	486	98
1404	gi9955912	Homo sapiens	GPVI mRNA for platelet glycoprotein VI-2, complete cds.	288	37
1404	AAB40232	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 46 SEQ ID NO:142.	326	40
1405	AAM06606	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 337.	50	34
1405	gi495989	Homo sapiens	Human rearranged IgH chain gene, VJ6 region, partial cds.	58	27
1405	AAM85487	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:13080.	56	32
1406	gi13377867	Gallus gallus	claudin-3	116	24
1406	gi15553371	Danio rerio	claudin c	112	26

Table 2A  
182

SEQ ID	Hit ID	Species	Description	S score	% Identity
1406	gi15553375	Danio rerio	claudin h	110	23
1407	AAAY78801	Homo sapiens	PROT- Hydrophobic domain containing protein clone HP00631 amino acid sequence.	701	100
1407	AAAY32204	Homo sapiens	INCY- Human receptor molecule (REC) encoded by Incyte clone 2132179.	701	100
1407	gi5231135	Homo sapiens	androgen induced protein (AIG-1) mRNA, complete cds.	695	99
1408	gi13543940	Homo sapiens	Similar to RIKEN cDNA 2610017G09 gene, clone MGC:12975 IMAGE:3347312, mRNA, complete cds.	2232	96
1408	AAB12138	Homo sapiens	PROT- Hydrophobic domain protein isolated from HT-1080 cells.	2221	96
1408	AAG81335	Homo sapiens	ZYMO Human AFP protein sequence SEQ ID NO:188.	2218	96
1410	gi10719608	Homo sapiens	IL-22 receptor (IL22R) mRNA, complete cds.	1699	100
1410	AAW97861	Homo sapiens	ZYMO Human cytokine receptor 11 (Zcytor11).	1699	100
1410	AAAY97045	Homo sapiens	MILL- Human TANGO 241.	1699	100
1411	gi14090278	Rattus norvegicus	TAT1	668	84
1411	gi458247	Homo sapiens	Human X-linked PEST-containing transporter (XPCT) mRNA, partial cds.	414	51
1411	gi458255	Homo sapiens	Human X-linked PEST-containing transporter (XPCT) gene, exon 6.	414	51
1412	gi4378057	Homo sapiens	organic anion transporter 1 (OAT1) mRNA, complete cds.	317	51
1412	gi4579725	Homo sapiens	mRNA for hOAT1-2, complete cds.	317	51
1412	gi5901645	Homo sapiens	organic anion transporter 1 (SLC22A6) mRNA, complete cds.	317	51
1413	gi3881524	Caenorhabditis elegans	ZK1067.4	714	41
1413	gi19322	Lycopersicon esculentum	glycine-rich protein	63	50
1413	gi2204081	Pinctada fucata	insoluble protein	93	52
1414	AAB43682	Homo sapiens	HUMA- Human cancer associated protein sequence SEQ ID NO:1127.	1522	100
1414	gi12654351	Homo sapiens	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3, clone MGC:5280 IMAGE:2984830, mRNA, complete cds.	1522	100
1414	gi12654961	Homo sapiens	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3, clone MGC:5556 IMAGE:3457151, mRNA, complete cds.	1522	100
1415	gi1764015	Ciona intestinalis	COS41.5	314	42
1415	AAM42167	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 7098.	284	30

Table 2A

183

SEQ ID	Hit ID	Species	Description	S score	% Identity
1415	AAE03438	Homo sapiens	HUMA- Human gene 12 encoded secreted protein HETHW90, SEQ ID NO: 121.	258	30
1416	gi13591714	Homo sapiens	immunoglobulin superfamily receptor translocation associated protein 2c (IRTA2) mRNA, complete cds, alternatively spliced.	1128	88
1416	gi15277746	Homo sapiens	Fc receptor-like protein 5 (FCRH5) mRNA, complete cds.	1128	88
1416	AAB82315	Homo sapiens	UYCO Human immunoglobulin receptor isoform IRTA2c.	1128	88
1417	AAV40386_aal	Homo sapiens	INCY- Human zinc binding protein ZB-2 encoding cDNA.	525	80
1417	gi13278762	Homo sapiens	ring finger protein 5, clone MGC:2407 IMAGE:2822537, mRNA, complete cds.	525	80
1417	gi13366064	Homo sapiens	mRNA for HsRma1, complete cds.	525	80
1418	gi3077703	Oryctolagus cuniculus	mitsugumin29	1336	93
1418	gi3461888	Mus musculus	mitsugumin29	1314	91
1418	AAU25436	Homo sapiens	INCY- Human mddt protein from clone LG:171377.1:2000MAY19.	1050	96
1419	gi13452508	Mus musculus	claudin 14	371	40
1419	gi12597447	Homo sapiens	claudin 14 (CLDN14) mRNA, complete cds.	370	40
1419	gi15082421	Homo sapiens	Similar to claudin 14, clone MGC:20195 IMAGE:4684949, mRNA, complete cds.	370	40
1420	AAG00539	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 4620.	173	49
1420	AAU20426	Homo sapiens	HUMA- Human secreted protein, Seq ID No 418.	163	47
1420	AAG75413	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:6177.	159	50
1421	gi14486155	Bos taurus	Rh type B glycoprotein	1703	86
1421	gi15718471	Homo sapiens	Rh type B glycoprotein (RHBG) gene, exons 9, and 10 and complete cds.	1448	80
1421	gi9858562	Homo sapiens	Rh type B glycoprotein (RHBG) mRNA, complete cds.	1448	80
1422	AAM00949	Homo sapiens	HYSE- Human bone marrow protein, SEQ ID NO: 425.	215	46
1422	AAE01852	Homo sapiens	HUMA- Human gene 11 encoded secreted protein fragment, SEQ ID NO:175.	174	40
1422	gi14209834	Mus musculus	ATP-binding cassette transporter sub-family A member 7	178	38
1423	gi12053628	Homo sapiens	mRNA for ribonuclease 7.	718	99
1423	AAY44192	Homo sapiens	INNO- Human keratinocyte-derived RNase-like protein.	718	99
1423	AAB10601	Homo sapiens	SCHD Human SAP-2 pre-protein.	718	99
1424	AAM83996	Homo sapiens	HUMA- Human immune/haematopoietic antigen SEQ ID NO:11589.	1153	99
1424	AAG02219	Homo sapiens	GEST Human secreted protein, SEQ ID NO: 6300.	483	99

Table 2A  
184

SEQ ID	Hit ID	Species	Description	S score	% Identity
1424	gi222902	Oncorhynchus keta	pituitary-specific transcription factor, Pit-1/GHF-1	98	32
1425	AAB82485	Homo sapiens	ZYMO Human secretin-like receptor Zgpr1.	499	92
1425	AAB82487	Homo sapiens	ZYMO Human secretin-like receptor Zgpr1 splice variant.	499	92
1425	AAE03382	Homo sapiens	HUMA- Human gene 5 encoded secreted protein HEOMX53, SEQ ID NO:40.	499	92
1426	gi6808374	Homo sapiens	mRNA; cDNA DKFZp434G0812 (from clone DKFZp434G0812); partial cds.	3211	100
1426	gi433383	Tripneustes gratilla	dynein heavy chain isotype 5A	241	26
1426	gi6706264	Leishmania major	dynein heavy chain	271	26
1428	gi13540300	Mus musculus	nucleolar protein C7B	427	34
1428	gi13561516	Mus musculus	nucleolar protein C7	426	34
1428	AAM25939	Homo sapiens	HYSE- Human protein sequence SEQ ID NO:1454.	156	90
1429	AAU27632	Homo sapiens	ZYMO Human protein AFP674535.	117	49
1429	AAM40391	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 3536.	117	49
1429	gi7022187	Homo sapiens	cDNA FLJ10261 fis, clone HEMBB1000975.	85	42
1430	AAAY07751	Homo sapiens	HUMA- Human secreted protein fragment encoded from gene 8.	293	100
1430	gi1142588	Trypanosoma brucei	CR3	84	43
1430	AAO05990	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 19882.	72	33
1431	AAM93525	Homo sapiens	HELI- Human polypeptide, SEQ ID NO: 3259.	246	36
1431	AAAY66693	Homo sapiens	GETH Membrane-bound protein PRO1004.	191	38
1431	AAB65216	Homo sapiens	GETH Human PRO1004 (UNQ488) protein sequence SEQ ID NO:227.	191	38
1432	AAB88388	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0131.	316	44
1432	AAB25719	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 6 SEQ ID NO:108.	84	73
1432	AAM62047	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 34152.	45	39
1433	AAB88388	Homo sapiens	HELI- Human membrane or secretory protein clone PSEC0131.	318	44
1433	AAB25719	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 6 SEQ ID NO:108.	86	66
1433	gi5629917	Homo sapiens	partial FLN2 gene for ABP-L, gamma filamin, exons 1 to 3.	77	34
1434	AAG75991	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:6755.	235	77



Table 2A

185

SEQ ID	Hit ID	Species	Description	S score	% Identity
1434	AAM78941	Homo sapiens	HYSE- Human protein SEQ ID NO 1603.	220	86
1434	ABB14802	Homo sapiens	HUMA- Human nervous system related polypeptide SEQ ID NO 3459.	58	40
1435	gi9621664	Homo sapiens	RHBDL gene for rhomboid-related protein.	1129	56
1435	gi14336709	Homo sapiens	16p13.3 sequence section 3 of 8.	1123	56
1435	gi3287191	Homo sapiens	mRNA for rhomboid-related protein, complete CDS.	1123	56
1436	gi7106866	Homo sapiens	HSPC238	238	56
1436	AAU15887	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 840.	238	56
1436	AAU16342	Homo sapiens	HUMA- Human novel secreted protein, Seq ID 1295.	238	56
1437	gi1418942	Mus musculus	semaphorin G	5660	93
1437	AAAY94990	Homo sapiens	ALPH- Human secreted protein vb21_1, SEQ ID NO:20.	5403	99
1437	gi2772584	Homo sapiens	semaphorin F homolog mRNA, complete cds.	3560	59
1439	gi11055322	Homo sapiens	vanilloid receptor-related osmotically activated channel (VROAC) mRNA, complete cds.	3324	100
1439	AAI66972_aa1	Homo sapiens	MILL- Human ion channel VR-5 cDNA sequence.	3319	99
1439	AAG65787	Homo sapiens	MILL- Human ion channel VR-5 protein sequence.	3318	99
1440	gi4155033	Helicobacter pylori J99	cag island protein	72	25
1440	AAU69567	Homo sapiens	PHAA Human G protein-coupled receptor from cDNA Seq-2643.	77	23
1440	gi13171062	turkey coronavirus	M protein	56	41
1441	gi16554186	Homo sapiens	cDNA FLJ25409 fis, clone TST03074.	644	100
1441	gi4235228	Mus musculus	leucine zipper-EF-hand containing transmembrane protein 1	497	59
1441	gi6599194	Homo sapiens	mRNA; cDNA DKFZp434C229 (from clone DKFZp434C229); partial cds.	493	63
1442	gi16359165	Homo sapiens	clone IMAGE:4645529, mRNA, partial cds.	1270	100
1442	ABB11242	Homo sapiens	HYSE- Human SLIT-2 homologue, SEQ ID NO:1612.	653	99
1442	AAB07469	Homo sapiens	ZYMO A human leucine-rich repeat protein designated Zlrr3.	451	36
1443	AAE03245	Homo sapiens	HUMA- Human gene 3 encoded secreted protein fragment, SEQ ID NO:95.	804	100
1443	AAE03244	Homo sapiens	HUMA- Human gene 3 encoded secreted protein fragment, SEQ ID NO:94.	799	100
1443	AAE03204	Homo sapiens	HUMA- Human gene 3 encoded secreted protein HNGNN78, SEQ ID NO:54.	617	100
1444	gi4062463	Escherichia coli	ABC transporter probable ATP-	697	91

Table 2A

186

SEQ ID	Hit ID	Species	Description	S score	% Identity
			binding subunit homolog		
1444	gi15487341	Escherichia coli	macrolide-specific ABC-type efflux carrier	697	91
1444	gi17743610	Agrobacterium tumefaciens str. C58 (Dupont)	ABC transporter, nucleotide binding/ATPase protein	366	51
1446	gi3168602	Homo sapiens	p160 mRNA, partial cds.	3755	89
1446	AAW31185	Homo sapiens	DAND Human p160 polypeptide 160.1.	3169	87
1446	gi3168604	Homo sapiens	proline and glutamic acid rich nuclear protein isoform mRNA, partial cds.	1641	99
1447	gi14042515	Homo sapiens	cDNA FLJ14761 fis, clone NT2RP3003302.	67	36
1447	AAG67254	Homo sapiens	HELI- Amino acid sequence of a human liver-associated gene.	67	36
1447	AAB94495	Homo sapiens	HELI- Human protein sequence SEQ ID NO:15188.	67	36
1448	AAB24058	Homo sapiens	GETH Human PRO290 protein sequence SEQ ID NO:7.	1972	100
1448	AAV66639	Homo sapiens	GETH Membrane-bound protein PRO290.	1972	100
1448	AAB65162	Homo sapiens	GETH Human PRO290 (UNQ253) protein sequence SEQ ID NO:33.	1972	100
1449	gi14043409	Homo sapiens	Similar to procollagen, type IV, alpha 3, clone MGC:11337 IMAGE:3953131, mRNA, complete cds.	89	75
1449	gi5420387	Leishmania major	proteophosphoglycan	86	24
1449	gi16117372	Macropodid herpesvirus 1	ICP4	102	27
1450	AAU27660	Homo sapiens	ZYMO Human protein AFP671052.	889	100
1450	AAG74151	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:4915.	748	100
1450	gi5670326	Homo sapiens	ICERE-1 mRNA, complete cds.	110	31
1451	AAV86519	Homo sapiens	HUMA- Human gene 71-encoded protein fragment, SEQ ID NO:434.	46	60
1451	AAM59183	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 31288.	61	39
1451	AAM71718	Homo sapiens	MOLE- Human bone marrow expressed probe encoded protein SEQ ID NO: 32024.	61	39
1452	gi12053219	Homo sapiens	mRNA; cDNA DKFZp434N1235 (from clone DKFZp434N1235); complete cds.	1031	93
1452	gi15559050	Ethmostigmus rubripes	ADP-ATP translocator	738	68
1452	gi339723	Homo sapiens	Human ADP/ATP translocase mRNA, 3' end, clone pHAT8.	693	68
1453	gi15025781	Clostridium acetobutylicum	Predicted membrane protein	136	32
1453	AAM95190	Homo sapiens	HUMA- Human reproductive system related antigen SEQ ID NO: 3848.	42	23

Table 2A  
187

SEQ ID	Hit ID	Species	Description	S score	% Identity
1453	AAM55650	Homo sapiens	MOLE- Human brain expressed single exon probe encoded protein SEQ ID NO: 27755.	61	31
1454	gi4929597	Homo sapiens	CGI-64 protein mRNA, complete cds.	1879	96
1454	gi6995987	Homo sapiens	mitochondrial carrier homolog 1 isoform a mRNA, complete cds; nuclear gene for mitochondrial product.	1818	99
1454	gi6995989	Homo sapiens	mitochondrial carrier homolog 1 isoform b (MTCH1) mRNA, partial cds; nuclear gene for mitochondrial product.	1483	99
1455	gi17131893	Nostoc sp. PCC 7120	WD-repeat protein	248	26
1455	gi886024	Thermomonospora curvata	PkwA	248	30
1455	gi17225210	Podospora anserina	beta transducin-like protein HET-D2Y	250	25
1456	AAB36840	Homo sapiens	ZYMO Human insulin receptor-related receptor protein with signal peptide.	6733	98
1456	gi186555	Homo sapiens	Human insulin receptor-related receptor (IRR) mRNA, 3' end.	6728	99
1456	AAB36836	Homo sapiens	ZYMO Human insulin receptor-related receptor protein.	6728	99
1457	gi6453436	Homo sapiens	mRNA; cDNA DKFZp586E041 (from clone DKFZp586E041); partial cds.	10115	100
1457	ABB11803	Homo sapiens	HYSE- Human GPI-122 homologue, SEQ ID NO:2173.	6423	99
1457	AAV50125	Homo sapiens	GEMY Human glycerophosphatidylinositol-anchored protein GPI-122.	6323	100
1458	AAU00023	Homo sapiens	BIOJ Human activated T-lymphocyte associated sequence 2, ATLAS-2.	3623	99
1458	AAE04546	Homo sapiens	INCY- Human G-protein coupled receptor-2 (GCREC-2) protein.	2570	79
1458	ABB11735	Homo sapiens	HYSE- Human vasopressin receptor homologue, SEQ ID NO:2105.	2546	100
1459	gi7021924	Homo sapiens	cDNA FLJ10081 fis, clone HEMBA1002018.	2742	100
1459	AAB92508	Homo sapiens	HELI- Human protein sequence SEQ ID NO:10631.	2742	100
1459	gi10435862	Homo sapiens	cDNA FLJ13751 fis, clone PLACE3000339, weakly similar to GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3).	2687	99
1460	AAM95163	Homo sapiens	HUMA- Human reproductive system related antigen SEQ ID NO: 3821.	75	33
1460	AAM06875	Homo sapiens	HYSE- Human foetal protein, SEQ ID NO: 1083.	68	45
1460	AAG76978	Homo sapiens	HUMA- Human colon cancer antigen protein SEQ ID NO:7742.	65	39

Table 2A  
188

SEQ ID	Hit ID	Species	Description	S score	% Identity
1461	gi1353365	Macaca fascicularis	MHC-G	84	30
1461	gi2655072	Pan troglodytes	MHC class I antigen HLA-H ortholog	72	28
1461	gi2655078	Pan paniscus	MHC class I antigen HLA-H ortholog	72	28
1462	gi11066090	Homo sapiens	matrix metalloprotease MMP-27 mRNA, complete cds.	1191	85
1462	AAE10410	Homo sapiens	SCHA/ Human full length matrix metalloproteinase-251 (MMP-251) protein.	1188	85
1462	AAAY90293	Homo sapiens	INCY- Human peptidase, HPEP-10 protein sequence.	1188	85
1463	gi601948	Drosophila melanogaster	Inscuteable	126	25
1463	gi1657962	Drosophila melanogaster	Nem	115	26
1463	gi15281684	Bacteriophage Mx8	p4	66	40
1464	AAB45378	Homo sapiens	HUMA- Human secreted protein sequence encoded by gene 38 SEQ ID NO:130.	434	87
1464	gi406058	Mus musculus	protein kinase	219	54
1464	gi13537204	Homo sapiens	mRNA for MAST205, complete cds.	216	54
1466	AAAY97293	Homo sapiens	INCY- Lipid associated protein (LIPAP) 3335404CD1.	1894	77
1466	AAB24231	Homo sapiens	INCY- Human vesicle associated protein 10 SEQ ID NO:10.	1116	48
1466	AAM39997	Homo sapiens	HYSE- Human polypeptide SEQ ID NO 3142.	1116	48
1467	gi13278855	Homo sapiens	calcium binding atopy-related autoantigen 1, clone MGC:2891 IMAGE:3009677, mRNA, complete cds.	919	96
1467	gi13278921	Homo sapiens	calcium binding atopy-related autoantigen 1, clone MGC:4521 IMAGE:3009677, mRNA, complete cds.	919	96
1467	AAB58329	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 667.	919	96
1468	gi12654401	Homo sapiens	Similar to RNA cyclase homolog, clone MGC:1390 IMAGE:3343468, mRNA, complete cds.	690	78
1468	gi10434565	Homo sapiens	cDNA FLJ12842 fis, clone NT2RP2003286, weakly similar to PROBABLE RNA 3'-TERMINAL PHOSPHATE CYCLASE (EC 6.5.1.4).	690	78
1468	AAB97250	Homo sapiens	HOMO RNA cyclase 41 protein.	690	78
1469	gi12053215	Homo sapiens	mRNA; cDNA DKFZp434K2435 (from clone DKFZp434K2435); complete cds.	273	98
1469	gi2633333	Bacillus subtilis	yhaJ	57	25
1469	gi15023682	Clostridium acetobutylicum	Glycosyltransferase involved in cell wall biogenesis	73	34

Table 2A

189

SEQ ID	Hit ID	Species	Description	S score	% Identity
1470	gi14290599	Homo sapiens	clone MGC:17624 IMAGE:3855543, mRNA, complete cds.	347	100
1470	gi65265	Xenopus laevis	a xenopus upstream binding factor	131	29
1470	gi65201	Xenopus laevis	RNA polymerase I transcription factor	128	29
1471	gi13182747	Homo sapiens	microsomal signal peptidase subunit mRNA, complete cds.	148	96
1471	AAW29660	Homo sapiens	Homo sapiens CH27_1 clone secreted protein.	148	96
1471	gi164084	Canis familiaris	signal peptidase 21 kDa subunit	141	90
1472	AAG03600	Homo sapiens	Human secreted protein, SEQ ID NO: 7681.	245	86
1472	gi7770239	Homo sapiens	PRO2831	139	71
1472	gi14026000	Mesorhizobium loti	transposase	64	46
1473	gi3065951	Homo sapiens	Notch3 (NOTCH3) gene, exon 33 and complete cds.	99	58
1473	gi2668592	Homo sapiens	Notch3 (NOTCH3) mRNA, complete cds.	99	58
1473	AAW49698	Homo sapiens	Human Notch3 protein.	99	58
1474	gi458938	Saccharomyces cerevisiae	Yhr186cp	156	58
1474	gi5921144	Schizosaccharom yces pombe	mip1	151	83
1474	gi6459542	Deinococcus radiodurans	serine/threonine protein kinase- related protein	96	31
1475	gi6562173	Homo sapiens	mRNA; cDNA DKFZp566H033 (from clone DKFZp566H033); partial cds.	295	91
1475	AAB38280	Homo sapiens	Human secreted protein sequence encoded by gene 20 SEQ ID NO:136.	289	96
1475	AAB94892	Homo sapiens	Human protein sequence SEQ ID NO:16234.	284	89
1476	gi7160973	Homo sapiens	mRNA for VNN3 protein.	1954	93
1476	gi6102996	Mus musculus	Vanin-3	1661	73
1476	gi6649540	Canis familiaris	TIFF66	1370	67
1477	gi7959741	Homo sapiens	PRO1051	134	76
1477	gi172903	Saccharomyces cerevisiae	transcription factor IIIA	113	26
1477	gi786305	Saccharomyces cerevisiae	Transcription factor TFIIIA (PIR accession number S20050)	113	26
1478	gi12652825	Homo sapiens	voltage-dependent anion channel 2, clone MGC:5237 IMAGE:2901130, mRNA, complete cds.	483	74
1478	gi15277577	Homo sapiens	voltage-dependent anion channel 2, clone MGC:21498 IMAGE:3875077, mRNA, complete cds.	483	74
1478	gi5114261	Homo sapiens	voltage-dependent anion channel isoform 2 (VDAC2) gene, exon 10 and complete cds.	483	74
1479	gi10435380	Homo sapiens	cDNA FLJ13381 fis, clone PLACE1001010.	133	58

Table 2A  
190

SEQ ID	Hit ID	Species	Description	S score	% Identity
1479	AAB94622	Homo sapiens	Human protein sequence SEQ ID NO:15476.	133	58
1479	AAB85361	Homo sapiens	Human phosphatase (PP) (clone ID 7472032CD1).	118	50
1480	gi12803725	Homo sapiens	APG5 (autophagy 5, S. cerevisiae)-like, clone MGC:3622 IMAGE:3609927, mRNA, complete cds.	204	100
1480	gi2995198	Homo sapiens	H.sapiens mRNA for apoptosis specific protein.	204	100
1480	gi12006864	Homo sapiens	apoptosis-related protein (APG5L) mRNA, complete cds, alternatively spliced.	204	100
1481	gi12619679	Conus arenatus	conotoxin scaffold VI/VII precursor	56	25
1481	gi2661493	Drosophila melanogaster	56F3.e	72	34
1481	gi2648663	Archaeoglobus fulgidus	DNA-directed RNA polymerase, subunit H (rpoH)	55	27
1482	gi12654623	Homo sapiens	RAB31, member RAS oncogene family, clone MGC:1258 IMAGE:3534853, mRNA, complete cds.	194	58
1482	gi1457954	Homo sapiens	Human small GTP-binding protein rab22b mRNA, complete cds.	194	58
1482	gi1388195	Homo sapiens	Human low-Mr GTP-binding protein (RAB31) mRNA, complete cds.	194	58
1483	gi7959778	Homo sapiens	PRO1546	163	65
1483	AAG02753	Homo sapiens	Human secreted protein, SEQ ID NO: 6834.	154	68
1483	AAB93922	Homo sapiens	Human protein sequence SEQ ID NO:13907.	137	59
1484	gi6979921	Drosophila melanogaster	RhoGTPase	160	32
1484	gi7263024	Drosophila melanogaster	G protein RhoBTB	160	32
1484	gi15291731	Drosophila melanogaster	LD24835p	160	32
1485	gi14150450	Rattus norvegicus	UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase T9	246	93
1485	gi3047203	Caenorhabditis elegans	GLY7	128	54
1485	gi304259	Bos taurus	UDP-GalNAc:polypeptide, N-acetylgalactosaminyltransferase	109	45
1486	AAB95830	Homo sapiens	Human protein sequence SEQ ID NO:18850.	202	76
1486	AAG03710	Homo sapiens	Human secreted protein, SEQ ID NO: 7791.	192	73
1486	AAG02922	Homo sapiens	Human secreted protein, SEQ ID NO: 7003.	182	69
1487	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	1244	84
1487	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:121.	1249	85
1487	gi339771	Homo sapiens	Human transposon L1.1 with a	1243	84

Table 2A  
191

SEQ ID	Hit ID	Species	Description	S score	% Identity
			base deletion relative to L1.2B resulting in a premature stop codon in the coding region.		
1488	gi13560707	Homo sapiens	group XIII secreted phospholipase A2 mRNA, complete cds.	383	98
1488	gi15824793	Homo sapiens	group XIII secreted phospholipase A2	383	98
1488	AAV27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	383	98
1489	gi2995442	Homo sapiens	mRNA for UDPGal:GlcNAc b1,4 galactosyltransferase.	996	98
1489	gi4520136	Homo sapiens	mRNA for beta-1,4-galactosyltransferase II, complete cds.	996	98
1489	gi3869131	Mus musculus	beta-1,4-galactosyltransferase II	946	94
1490	AAM06551	Homo sapiens	Human foetal protein, SEQ ID NO: 282.	140	49
1490	AAB51718	Homo sapiens	Human secreted protein sequence encoded by gene 45 SEQ ID NO:158.	129	58
1490	gi1196431	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	124	52
1491	gi12052884	Homo sapiens	mRNA; cDNA DKFZp564C2478 (from clone DKFZp564C2478); complete cds.	487	98
1491	gi7023332	Homo sapiens	cDNA FLJ10961 fis, clone PLACE1000588, highly similar to INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1.	487	98
1491	AAB93371	Homo sapiens	Human protein sequence SEQ ID NO:12521.	487	98
1492	AAG00392	Homo sapiens	Human secreted protein, SEQ ID NO: 4473.	150	81
1492	AAB54106	Homo sapiens	Human pancreatic cancer antigen protein sequence SEQ ID NO:558.	136	71
1492	gi914110	Streptococcus pyogenes	EmmL15	107	20
1493	gi6855513	Gallus gallus	syndesmos	554	60
1493	gi13623247	Homo sapiens	Similar to RIKEN cDNA 1110001K21 gene, clone MGC:11275 IMAGE:3944355, mRNA, complete cds.	534	58
1493	gi12544542	Corynebacterium glutamicum	RXA02115	99	28
1494	gi12082725	Mus musculus	B cell phosphoinositide 3-kinase adaptor	335	62
1494	gi12082811	Gallus gallus	B cell phosphoinositide 3-kinase adaptor	211	48
1494	gi330842	Equine herpesvirus 1	myristylated virion protein	61	33
1495	AAB43811	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1256.	273	100
1495	AAB64482	Homo sapiens	Human secreted protein sequence encoded by gene 13 SEQ ID NO:120.	273	100
1495	gi5106795	Homo sapiens	sec61 homolog mRNA, complete	273	100

Table 2A  
192

SEQ ID	Hit ID	Species	Description	S score	% Identity
			cds.		
1496	AAE01786	Homo sapiens	Human gene 17 encoded secreted protein HWBEM18, SEQ ID NO:107.	4730	97
1496	gi6650678	Mus musculus	nuclear pore membrane glycoprotein POM210	4025	76
1496	gi56463	Rattus norvegicus	gp210 (AA 1-1886)	4000	79
1497	gi185996	Homo sapiens	Human Ig germline kappa L-chain V-region gene (HK166), V-kappa-1.	607	100
1497	AAR38651	Homo sapiens	Human V-kappa fragment encoded by clone vk65.15.	607	100
1497	AAR62931	Homo sapiens	Human V-kappa vk65.15 region.	607	100
1498	gi431857	Homo sapiens	H.sapiens mRNA for delta 4-3-oxosteroid 5 beta-reductase.	460	76
1498	gi11640835	Homo sapiens	5-beta steroid reductase (SRD5B1) gene, exon 9 and complete cds.	460	76
1498	gi5689216	Oryctolagus cuniculus	delta4-3-oxosteroid 5beta-reductase	442	70
1499	gi1752736	Saccharomyces cerevisiae	gene required for phosphoylation of oligosaccharides/ has high homology with YJR061w	210	47
1499	AAB53977	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:1517.	178	71
1499	AAB27956	Homo sapiens	Human secreted protein SEQ ID NO: 110.	177	65
1500	AAB93159	Homo sapiens	Human protein sequence SEQ ID NO:12081.	2209	64
1500	AAB58796	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 504.	1455	65
1500	gi13377567	Vibrio cholerae	accessory colonization factor AcfD	105	26
1501	AAB95655	Homo sapiens	Human protein sequence SEQ ID NO:18417.	186	78
1501	AAB95596	Homo sapiens	Human protein sequence SEQ ID NO:18279.	172	80
1501	gi1196433	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	182	78
1502	AAB58202	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 540.	427	100
1502	gi4572328	Homo sapiens	caveolin-1 gene, exon 3 and complete cds.	345	100
1502	gi6599075	Homo sapiens	caveolin-1/-2 locus, Contig1, D7S522, genes CAV2 (exons 1, 2a, and 2b), CAV1 (exons 1 and 2).	345	100
1503	gi199584	Mus musculus	MHox	405	96
1503	gi51362	Mus musculus	DNA-binding protein	405	96
1503	gi1836044	Rattus sp.	rHox protein	405	96
1504	gi8163762	Homo sapiens	membrane cofactor protein CD46 variant (MCP) mRNA, partial cds.	639	83
1504	AAB58394	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 732.	635	82
1504	AAG75528	Homo sapiens	Human colon cancer antigen	635	82



Table 2A

193

SEQ ID	Hit ID	Species	Description	S score	% Identity
			protein SEQ ID NO:6292.		
1505	gi15990400	Homo sapiens	clone IMAGE:3954884, mRNA, partial cds.	872	83
1505	gi3523113	Homo sapiens	prostate-specific transglutaminase (TGM4) gene, alternative spliced variant, exon 2 and partial cds.	92	46
1505	AAB54389	Homo sapiens	Human pancreatic cancer antigen protein sequence SEQ ID NO:841.	84	56
1506	AAB94891	Homo sapiens	Human protein sequence SEQ ID NO:16231.	214	65
1506	gi1196431	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	197	66
1506	AAB38280	Homo sapiens	Human secreted protein sequence encoded by gene 20 SEQ ID NO:136.	196	69
1507	gi1184173	Homo sapiens	Human nucleoporin 98 (NUP98) mRNA, complete cds.	1058	100
1507	gi11414896	Homo sapiens	NUP98 mRNA for nucleoporin, complete cds.	1058	100
1507	gi4545101	Homo sapiens	cell-line HeLa Nup98-Nup96 precursor splice variant 1 mRNA, complete cds.	1060	88
1508	gi5106521	Homo sapiens	K-Cl cotransporter KCC4 mRNA, complete cds.	1062	96
1508	gi10440500	Homo sapiens	mRNA for FLJ00098 protein, partial cds.	1062	96
1508	gi10440514	Homo sapiens	mRNA for FLJ00105 protein, partial cds.	1062	96
1509	gi6691968	Homo sapiens	Human DNA sequence from clone RP1-148M19 on chromosome Xp11.22-11.3 Contains ZNF741 (zinc finger protein), a ribosomal protein L23a pseudogene, STSs and GSSs, complete sequence.	293	81
1509	gi4096339	Homo sapiens	Human zinc finger protein (ZNF741) mRNA, complete cds.	293	81
1509	AAB21033	Homo sapiens	Human nucleic acid-binding protein, NuABP-37.	293	81
1510	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	426	57
1510	gi5052951	Homo sapiens	LINE1 element inserted in B-globin gene intron 2.	425	57
1510	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:121.	426	57
1511	gi16041769	Homo sapiens	clone MGC:23189 IMAGE:4854518, mRNA, complete cds.	313	55
1511	gi4454678	Homo sapiens	zinc finger protein 4	308	55
1511	gi186774	Homo sapiens	Human Kruppel related zinc finger protein (HTF10) mRNA, complete cds.	321	55
1512	gi14027838	Mesorhizobium loti	transcriptional regulator	91	26
1513	gi4886463	Homo sapiens	mRNA; cDNA DKFZp586G1219 (from clone DKFZp586G1219); partial cds.	1039	100

Table 2A

194

SEQ ID	Hit ID	Species	Description	S score	% Identity
1513	gi4337460	Homo sapiens	neuroblastoma-amplified protein mRNA, complete cds.	1039	100
1513	AAB38417	Homo sapiens	Fragment of human secreted protein encoded by gene 5 clone HCGMF16.	1039	100
1514	gi1621611	Homo sapiens	Human TRAF family member-associated NF-kB activator TANK mRNA, complete cds.	1667	76
1514	gi1518018	Homo sapiens	Human TRAF-interacting protein I-TRAF mRNA, complete cds.	1655	77
1514	AAW27163	Homo sapiens	Human TRAF inhibitor protein I-TRAF.	1655	77
1515	gi12957169	Mus musculus	synaptotagmin-like protein 3-b	142	60
1515	gi13647085	Mus musculus	synaptotagmin-like protein 3-a delta 3S-II	142	60
1515	gi13647079	Mus musculus	synaptotagmin-like protein 3-a + 3S-I	142	60
1516	gi6688199	Homo sapiens	mRNA for AMP-activated protein kinase gamma2 subunit (AMPK gamma2 gene).	1055	93
1516	gi5931569	Homo sapiens	mRNA for H91620p, complete cds.	1055	93
1516	gi12642942	Homo sapiens	AMP-activated protein kinase gamma subunit (PRKAG2) mRNA, complete cds.	1055	93
1517	gi6807587	Homo sapiens	Novel human gene mapping to chromosome 1.	2360	100
1517	gi1769491	Homo sapiens	Human kruppel-related zinc finger protein (ZNF184) mRNA, partial cds.	1135	49
1517	gi186774	Homo sapiens	Human Kruppel related zinc finger protein (HTF10) mRNA, complete cds.	904	41
1518	gi9956065	Homo sapiens	clone CDABP0092 mRNA sequence.	602	100
1518	gi4038733	Homo sapiens	mRNA for beta 2-microglobulin, complete cds.	602	100
1518	gi5725512	Homo sapiens	beta-2 microglobulin gene, complete cds.	602	100
1519	gi187177	Homo sapiens	Human lamin-like protein in HindIII repetitive element derived DNA, 3' end.	216	57
1519	gi339771	Homo sapiens	Human transposon L1.1 with a base deletion relative to L1.2B resulting in a premature stop codon in the coding region.	223	57
1519	gi5070622	Homo sapiens	retrotransposon L1 insertion in X-linked retinitis pigmentosa locus, complete sequence.	223	57
1520	AAE03963	Homo sapiens	Human gene 17 encoded secreted protein fragment, SEQ ID NO:142.	110	47
1520	gi4097459	Elephantulus edwardii	reverse transcriptase	109	47
1520	AAB94930	Homo sapiens	Human protein sequence SEQ ID NO:16405.	109	46
1521	gi10934047	Mus musculus	Scot-t1	1035	77
1521	gi10934052	Mus musculus	Scot-t2	1035	77

Table 2A

195

SEQ ID	Hit ID	Species	Description	S score	% Identity
1521	gi164423	Sus scrofa	succinyl-CoA:alpha-ketoacid coenzyme A transferase	1024	81
1522	gi12003128	Eremothecium gossypii	Yer154p	108	31
1522	gi5052482	Drosophila melanogaster	BcDNA.GH02220	102	29
1522	gi6227006	Arabidopsis thaliana	F16G16.8	94	50
1523	gi3419880	Homo sapiens	mRNA for MDC/ADAM11, complete cds.	2989	100
1523	gi836683	Homo sapiens	Human metalloprotease/disintegrin-like (MDC) gene, partial cds.	2989	100
1523	AAR75352	Homo sapiens	Human fetal brain MDC protein.	2984	99
1524	gi1109782	Homo sapiens	Human protein-tyrosine phosphatase mRNA, complete cds.	2518	97
1524	gi1781037	Mus musculus	neuronal tyrosine threonine phosphatase 1	1996	87
1524	AAB66436	Homo sapiens	Human MAP-kinase phosphatase hVH5.	883	99
1525	gi757911	Homo sapiens	H.sapiens mRNA for A2b adenosine receptor.	442	100
1525	gi178150	Homo sapiens	Human adenosine A2b receptor (ADORA2) mRNA, complete cds.	442	100
1525	AAR41526	Homo sapiens	Human A2b adenosine receptor.	442	100
1526	gi13540160	Homo sapiens	TCF12-TEC fusion protein mRNA, partial cds.	335	79
1526	gi183930	Homo sapiens	Human HEB helix-loop-helix protein (HEB) mRNA, complete cds.	337	75
1526	gi184448	Homo sapiens	transcription factor (HTF4) mRNA, complete cds.	337	75
1527	AAB43940	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1385.	741	95
1527	gi14602778	Homo sapiens	spermine synthase, clone MGC:2071 IMAGE:3506022, mRNA, complete cds.	690	95
1527	gi2198557	Homo sapiens	spermidine aminopropyltransferase mRNA, complete cds.	690	95
1528	gi13436152	Homo sapiens	reticulocalbin 2, EF-hand calcium binding domain, clone MGC:1650 IMAGE:3505241, mRNA, complete cds.	674	81
1528	gi469885	Homo sapiens	H.sapiens ERC-55 mRNA.	674	81
1528	AAW21949	Homo sapiens	E6-binding protein E6-BPSD7.	674	81
1529	gi10440331	Homo sapiens	cDNA: FLJ23591 fis, clone LNG14729.	606	100
1529	AAB84327	Homo sapiens	Amino acid sequence of a human lyase and associated protein HLYAP-2.	606	100
1529	gi15150358	Mus musculus	UDP-glucuronic acid decarboxylase	605	99
1530	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	3094	92
1530	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID	3094	92

Table 2A  
196

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NO:121.		
1530	gi5052951	Homo sapiens	LINE1 element inserted in B-globin gene intron 2.	3083	91
1531	gi5596433	Homo sapiens	candidate tumor suppressor protein NOC2 (NOC2) mRNA, complete cds.	330	87
1531	gi7020579	Homo sapiens	cDNA FLJ20462 fis, clone KAT06107.	330	87
1531	gi13477353	Homo sapiens	Similar to rabphilin 3A-like (without C2 domains), clone MGC:3453 IMAGE:3529317, mRNA, complete cds.	329	87
1532	gi1531645	Rattus norvegicus	C2-HC type zinc finger protein r-Myt3	1914	67
1532	gi2914751	Rattus norvegicus	neural zinc finger factor 3; NZF-3	1909	66
1532	gi1531653	Xenopus laevis	C2-HC type zinc finger protein X-Myt1	300	42
1533	gi12805043	Homo sapiens	clone IMAGE:3461487, mRNA, partial cds.	364	41
1533	AAG81328	Homo sapiens	Human AFP protein sequence SEQ ID NO:174.	356	41
1533	gi2226004	Homo sapiens	Human Tigger1 transposable element, complete consensus sequence.	441	71
1534	gi10436783	Homo sapiens	cDNA FLJ14341 fis, clone THYRO1000343, weakly similar to ATROPHIN-1.	3184	97
1534	AAB95860	Homo sapiens	Human protein sequence SEQ ID NO:18924.	3184	97
1534	gi12802159	Homo sapiens	SH3-SAM adaptor protein (HACS1) mRNA, complete cds.	515	46
1535	gi15928572	Mus musculus	Similar to leucine rich repeat (in FLII) interacting protein 2	103	66
1535	gi14091821	Oryza sativa	Putative protein with region similar to cyclin-dependent kinase like proteins	98	38
1535	gi7020214	Homo sapiens	cDNA FLJ20248 fis, clone COLF6543.	97	63
1536	AAG02639	Homo sapiens	Human secreted protein, SEQ ID NO: 6720.	160	71
1536	AAG02753	Homo sapiens	Human secreted protein, SEQ ID NO: 6834.	141	60
1536	gi7959778	Homo sapiens	PRO1546	140	60
1537	gi7022610	Homo sapiens	cDNA FLJ10521 fis, clone NT2RP2000841.	184	39
1537	AAB92909	Homo sapiens	Human protein sequence SEQ ID NO:11539.	184	39
1537	gi11527193	Drosophila melanogaster	sunday driver	112	41
1538	gi12654055	Homo sapiens	clone IMAGE:3455871, mRNA, partial cds.	849	85
1538	gi3133291	Homo sapiens	mitogen activated protein kinase activated protein kinase gene, complete cds.	844	85
1538	gi2911813	Mus musculus	mitogen-activated protein kinase-activated protein kinase	841	84

Table 2A  
197

SEQ ID	Hit ID	Species	Description	S score	% Identity
1539	gi11761810	Arabidopsis thaliana	glutathione dependent dehydroascorbate reductase precursor	60	32
1540	gi13162677	Homo sapiens	GLUT4 enhancer factor mRNA, complete cds.	1399	99
1540	gi12655101	Homo sapiens	clone IMAGE:3140406, mRNA, partial cds.	1399	99
1540	AAB58934	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 642.	1395	99
1541	gi30058	Homo sapiens	Human mRNA for pro-alpha-1 type 3 collagen.	1606	98
1541	gi16197601	Homo sapiens	type III procollagen alpha 1 chain (COL3A1) gene, exon 1 and complete cds.	1606	98
1541	AAW12842	Homo sapiens	Truncated pro-alpha1(III) chain.	1606	98
1542	AAE01436	Homo sapiens	Human gene 1 encoded secreted protein HWLFJ10, SEQ ID NO:91.	1622	99
1542	AAE01464	Homo sapiens	Human gene 1 encoded secreted protein HWLFJ10, SEQ ID NO:119.	1618	98
1542	AAE01515	Homo sapiens	Human gene 1 encoded secreted protein fragment, SEQ ID NO:172.	1618	98
1543	gi186043	Homo sapiens	immunoglobulin light chain variable region (IGL@) mRNA, partial cds.	547	86
1543	gi219886	Homo sapiens	Human Ig kappa light chain gene, V- and J-region.	543	89
1543	gi33248	Homo sapiens	H.sapiens gene for Ig kappa light chain variable region '012'.	540	89
1544	gi7673618	Mus musculus	ubiquitin specific protease	1972	73
1544	gi7328168	Homo sapiens	mRNA; cDNA DKFZp434K1822 (from clone DKFZp434K1822); partial cds.	1004	74
1544	gi5823525	Drosophila melanogaster	ubiquitin-specific protease nonstop	957	41
1545	gi6693836	Rattus norvegicus	SNIP-b	3975	85
1545	gi6693834	Rattus norvegicus	SNIP-a	3975	85
1545	gi3098418	Mus musculus	P140	3699	75
1546	gi179433	Homo sapiens	Human biglycan (BGN) gene, exon 8.	2032	95
1546	gi12803217	Homo sapiens	biglycan, clone MGC:2298 IMAGE:3162633, mRNA, complete cds.	1512	96
1546	gi13279002	Homo sapiens	biglycan, clone MGC:10461 IMAGE:3503374, mRNA, complete cds.	1512	96
1547	gi14718648	Homo sapiens	allantoicase mRNA, partial cds.	777	95
1547	gi9255889	Mus musculus	allantoicase	1039	58
1547	gi4929823	Xenopus laevis	allantoicase	681	41
1548	gi12653161	Homo sapiens	ribosomal protein L35, clone MGC:8582 IMAGE:2960987, mRNA, complete cds.	203	78

Table 2A  
198

SEQ ID	Hit ID	Species	Description	S score	% Identity
1548	gi15012043	Homo sapiens	Similar to ribosomal protein L35, clone MGC:13488 IMAGE:4251487, mRNA, complete cds.	203	78
1548	gi562074	Homo sapiens	Human ribosomal protein L35 mRNA, complete cds.	203	78
1549	gi1903236	Mus musculus	capping protein beta 3 subunit	156	100
1549	gi595257	Homo sapiens	Human F-actin capping protein beta subunit mRNA, complete cds.	156	100
1549	gi500749	Mus musculus	capping protein beta subunit, isoform 2	156	100
1550	gi15278186	Homo sapiens	MAGI-1A mRNA, complete cds, alternatively spliced.	526	91
1550	gi3370998	Homo sapiens	mRNA for BAI1-associated protein 1, complete cds.	526	91
1550	gi15278182	Homo sapiens	MAGI-1B alpha beta mRNA, complete cds, alternatively spliced.	526	91
1551	gi12654299	Homo sapiens	clone IMAGE:3447394, mRNA, partial cds.	1945	99
1551	gi11095188	Homo sapiens	dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.	1357	67
1551	AAB47187	Homo sapiens	Human DPP8.	1357	67
1552	gi191012	Cricetus cricetus	ornithine decarboxylase	153	38
1552	gi49440	Cricetus cricetus	ornithine decarboxylase (AA 1-455)	153	38
1552	gi9858179	Danio rerio	ornithine decarboxylase	149	44
1553	gi12053087	Homo sapiens	mRNA; cDNA DKFZp434B0819 (from clone DKFZp434B0819); complete cds.	168	100
1553	gi16118555	Homo sapiens	ELMO1 mRNA, complete cds.	168	100
1553	gi16118551	Mus musculus	ELMO1	168	100
1554	AAY91640	Homo sapiens	Human secreted protein sequence encoded by gene 34 SEQ ID NO:313.	238	81
1554	AAB93352	Homo sapiens	Human protein sequence SEQ ID NO:12476.	238	81
1554	AAY91484	Homo sapiens	Human secreted protein sequence encoded by gene 34 SEQ ID NO:157.	233	100
1555	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	726	87
1555	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:121.	726	87
1555	gi1196433	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	727	87
1556	gi4530437	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240 mRNA, complete cds.	1125	90
1556	gi7109237	Drosophila melanogaster	TRAP240	213	44
1556	gi7230590	Drosophila melanogaster	Pap/DTRAP240	213	44
1557	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID	2330	89

Table 2A  
199

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NO:121.		
1557	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	2325	89
1557	gi5052951	Homo sapiens	LINE1 element inserted in B-globin gene intron 2.	2309	88
1558	gi11066463	Rattus norvegicus	RhoGEF glutamate transport modulator GTRAP48	3450	72
1558	gi7110160	Homo sapiens	guanine nucleotide exchange factor (LARG) mRNA, complete cds.	1063	50
1558	AAW64468	Homo sapiens	Human secreted protein from clone CW420_2.	1063	50
1559	gi10440888	Morone saxatilis	myosin heavy chain FM3A	456	51
1559	gi15982970	Danio rerio	myosin IIIA	476	55
1559	gi7958618	Homo sapiens	class III myosin (MYO3A) mRNA, complete cds, alternatively spliced.	456	51
1560	AAB93846	Homo sapiens	Human protein sequence SEQ ID NO:13688.	1112	66
1560	gi3834629	Mus musculus	diaphanous-related formin; p134 mDia2	565	28
1560	gi3171906	Homo sapiens	mRNA for dia-156 protein.	559	30
1561	AAB59019	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 727.	162	47
1561	AAG74843	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5607.	88	52
1561	AAB95751	Homo sapiens	Human protein sequence SEQ ID NO:18660.	85	30
1562	gi13195147	Mus musculus	HCH	1228	87
1562	AAW03515	Homo sapiens	Human DOCK180 protein.	1308	56
1562	gi1339910	Homo sapiens	Human DOCK180 protein mRNA, complete cds.	1304	56
1563	AAB52017	Homo sapiens	Human secreted protein sequence encoded by gene 6 SEQ ID NO:66.	266	100
1563	gi6449393	Callinectes sapidus	copper-specific metallothionein CuMT-II	57	37
1563	gi6010621	Human calicivirus strain BAV/2.1/98/DE U	RNA-dependent RNA polymerase	51	36
1564	gi540073	Homo sapiens	Human agouti gene, exon 3 and complete cds.	362	97
1564	gi608648	Homo sapiens	agouti signalling protein (ASP) gene, complete cds.	356	96
1564	AAW10102	Homo sapiens	Human agouti signalling protein.	352	84
1565	gi2463646	Homo sapiens	Human 3-hydroxy-3-methylglutaryl CoA synthase gene, exon 9 and complete cds.	1277	100
1565	gi619877	Homo sapiens	H.sapiens mRNA for 3-hydroxy-3-methylglutaryl coenzyme A synthase.	1277	100
1565	gi15928472	Mus musculus	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2	1155	88
1566	gi5912057	Homo sapiens	mRNA; cDNA DKFZp434B172 (from clone DKFZp434B172); partial cds.	350	98

Table 2A  
200

SEQ ID	Hit ID	Species	Description	S score	% Identity
1566	gi7578787	Homo sapiens	AD021 protein (AD021) mRNA, complete cds.	157	41
1566	AAB57036	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1614.	69	33
1567	gi2226005	Homo sapiens	Human Tigger1 transposable element, complete consensus sequence.	269	53
1567	gi7634783	Homo sapiens	HDCMB45P mRNA, partial cds.	210	43
1567	AAB39252	Homo sapiens	Human secreted protein sequence encoded by gene 12 SEQ ID NO:132.	126	61
1568	gi12653985	Homo sapiens	glycine cleavage system protein H (aminomethyl carrier), clone MGC:5190 IMAGE:3451361, mRNA, complete cds.	846	91
1568	gi184348	Homo sapiens	Human H-protein mRNA, complete cds.	846	91
1568	gi219671	Homo sapiens	mRNA for hydrogen carrier protein, a component of an enzyme complex, glycine synthase (EC 2.1.2.10).	846	91
1569	gi388109	Enterococcus faecalis	regulatory protein	64	47
1569	AAY25744	Homo sapiens	Human secreted protein encoded from gene 34.	56	50
1570	gi532505	Homo sapiens	Human bile acid CoA: Amino acid N-acyltransferase mRNA, complete cds.	2049	99
1570	gi15215152	Mus musculus	Similar to bile acid-Coenzyme A dehydrogenase: amino acid n-acyltransferase	1412	68
1570	gi604902	Rattus norvegicus	Kan-1	1394	68
1571	AAG89290	Homo sapiens	Human secreted protein, SEQ ID NO: 410.	198	97
1571	gi15487674	Homo sapiens	OSBP-related protein 1 mRNA, complete cds.	204	100
1571	AAB38248	Homo sapiens	Human secreted protein sequence encoded by gene 46 SEQ ID NO:104.	62	51
1572	gi4263739	Homo sapiens	BAC clone GS1-489L14 from 7p14-p12, complete sequence.	594	68
1572	gi6102812	Homo sapiens	GLI3 gene for GLI3 protein.	594	68
1572	gi183248	Homo sapiens	Human DNA-binding protein (GLI3) mRNA, complete cds.	594	68
1573	gi1019435	Trypanosoma cruzi	mucin-like protein	133	37
1573	gi1019433	Trypanosoma cruzi	mucin-like protein; Method: conceptual translation supplied by author	114	38
1573	gi1280434	Drosophila melanogaster	hemomucin	127	33
1574	AAB76873	Homo sapiens	Human lung tumour protein related protein sequence SEQ ID NO:798.	141	37
1574	gi6015472	Hylobates muelleri	dopamine receptor D4	133	34



Table 2A  
201

SEQ ID	Hit ID	Species	Description	S score	% Identity
1574	gi13421134	Caulobacter crescentus	translation initiation factor IF-2	155	34
1575	AAB74709	Homo sapiens	Human membrane associated protein MEMAP-15.	1294	100
1575	gi14495648	Homo sapiens	clone MGC:15606 IMAGE:3163718, mRNA, complete cds.	334	36
1575	gi15721997	Homo sapiens	zonadhesin (ZAN) gene, complete cds, alternatively spliced.	173	20
1576	gi1181628	Homo sapiens	H.sapiens APXL mRNA.	1624	78
1576	gi1773381	Homo sapiens	chromosome X clone U177G4, U152H5, U168D5, 174A6, U172D6, and U186B3 from Xp22, complete sequence.	1624	78
1576	AAW58988	Homo sapiens	Homo sapiens fetal kidney clone BD335_14 encoded protein.	317	71
1577	gi6979943	Homo sapiens	type 1 tumor necrosis factor receptor shedding aminopeptidase regulator mRNA, complete cds.	1499	86
1577	gi6381989	Homo sapiens	adipocyte-derived leucine aminopeptidase mRNA, complete cds.	1492	86
1577	gi6642987	Homo sapiens	aminopeptidase PILS (APPILS) mRNA, complete cds.	1492	86
1578	gi14626461	Rhizobium leguminosarum	HupE	82	28
1578	gi897	Canis familiaris	endothelin-2	64	36
1579	gi7657864	Homo sapiens	BAC clone RP11-236P2 from 2, complete sequence.	4025	98
1579	gi1374698	Homo sapiens	mRNA for nuclear protein, NP220, complete cds.	4014	98
1579	AAV07032	Homo sapiens	Breast cancer associated antigen precursor sequence.	4014	98
1580	gi551065	Mus musculus	protease-nexin 1	995	83
1580	gi14715029	Mus musculus	serine (or cysteine) proteinase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	995	83
1580	gi412265	Rattus norvegicus	glia-derived neurite-promoting factor (GdNPF)	994	82
1581	gi499184	Felis catus	neuronal protein	305	93
1581	gi10433455	Homo sapiens	cDNA FLJ12066 fis, clone HEMBB1002266, moderately similar to NEURONAL PROTEIN.	286	65
1581	AAB95041	Homo sapiens	Human protein sequence SEQ ID NO:16804.	286	65
1582	AAV13385	Homo sapiens	Amino acid sequence of protein PRO293.	3183	99
1582	AAB80253	Homo sapiens	Human PRO293 protein.	3183	99
1582	AAB33472	Homo sapiens	Human PRO1338 protein UNQ693 SEQ ID NO:279.	1676	55
1583	gi12803185	Homo sapiens	nucleophosmin (nucleolar phosphoprotein B23, numatrin), clone MGC:8463 IMAGE:2821577, mRNA, complete cds.	268	75

Table 2A  
202

SEQ ID	Hit ID	Species	Description	S score	% Identity
1583	gi14250152	Homo sapiens	nucleophosmin (nucleolar phosphoprotein B23, numatrin), clone MGC:14826 IMAGE:4276604, mRNA, complete cds.	268	75
1583	gi15214852	Homo sapiens	nucleophosmin (nucleolar phosphoprotein B23, numatrin), clone MGC:13433 IMAGE:4097025, mRNA, complete cds.	268	75
1584	gi4406691	Homo sapiens	clone 24922 mRNA sequence, complete cds.	697	100
1584	gi7023544	Homo sapiens	cDNA FLJ11094 fis, clone PLACE1005373, weakly similar to TRNA PSEUDOURIDINE SYNTHASE B (EC 4.2.1.70).	697	100
1584	AAB93492	Homo sapiens	Human protein sequence SEQ ID NO:12796.	697	100
1585	gi155999	Bombyx mori	silk fibroin	147	39
1585	gi930003	Bombyx mori	silk fibroin (AA 37 - 252)	147	39
1585	gi765323	Bombyx mori	silk fibroin heavy chain	158	39
1586	gi12002682	Homo sapiens	FERM-containing protein (CG1) mRNA, alternative splice product, complete cds.	1282	89
1586	gi7669988	Homo sapiens	mRNA; cDNA DKFZp761N1814 (from clone DKFZp761N1814).	1090	86
1586	AAB12318	Homo sapiens	Human secreted protein encoded by gene 18 clone HE2FL70.	189	100
1587	gi5019618	Homo sapiens	HFB30 mRNA, complete cds.	1005	90
1587	gi4530066	Homo sapiens	androgen receptor associated protein 54 (ARA54) mRNA, complete cds.	1005	90
1587	AAY78418	Homo sapiens	Human androgen receptor coactivator ARA54 SEQ ID NO:2.	1005	90
1588	AAG73501	Homo sapiens	Human secreted protein fragment, SEQ ID NO:277.	285	94
1588	gi1196432	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	285	94
1588	gi339771	Homo sapiens	Human transposon L1.1 with a base deletion relative to L1.2B resulting in a premature stop codon in the coding region.	285	94
1589	AAB94900	Homo sapiens	Human protein sequence SEQ ID NO:16288.	287	67
1589	gi1196433	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	289	70
1589	gi1916229	Homo sapiens	Human line-1 reverse transcriptase gene, partial cds, and granulocyte chemotactic protein-2 (GCP-2) gene, complete cds.	278	70
1590	gi6562173	Homo sapiens	mRNA; cDNA DKFZp566H033 (from clone DKFZp566H033); partial cds.	227	88
1590	gi1335199	Homo sapiens	Human KpnI repetitive sequence (T-betaG41) 3kb downstream of beta-globin gene.	227	86
1590	AAB38280	Homo sapiens	Human secreted protein sequence	226	88

Table 2A  
203

SEQ ID	Hit ID	Species	Description	S score	% Identity
			encoded by gene 20 SEQ ID NO:136.		
1591	gi10436007	Homo sapiens	cDNA FLJ13859 fis, clone THYRO1001033, weakly similar to TRANSFORMATION-SENSITIVE PROTEIN IEF SSP 3521.	1794	80
1591	AAB95593	Homo sapiens	Human protein sequence SEQ ID NO:18273.	1794	80
1591	gi7020708	Homo sapiens	cDNA FLJ20535 fis, clone KAT11013.	1789	79
1592	gi6599260	Homo sapiens	mRNA; cDNA DKFZp434P1721 (from clone DKFZp434P1721); partial cds.	1184	76
1592	gi10567164	Homo sapiens	GASC-1 mRNA, complete cds.	561	59
1592	AAR66461	Homo sapiens	AF-17 protein, N-terminal region with similarity to peregrin.	181	34
1593	gi182020	Homo sapiens	Human elastin gene, exon 1.	2405	89
1593	AAB08630	Homo sapiens	Amino acid sequence of a human elastin polypeptide.	2392	88
1593	gi182062	Homo sapiens	Human elastin mRNA, complete cds.	2386	88
1594	gi14424570	Homo sapiens	clone MGC:16614 IMAGE:4111344, mRNA, complete cds.	349	88
1594	gi1196433	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	349	89
1594	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	349	89
1595	gi291854	Homo sapiens	aminopeptidase A mRNA, complete cds.	934	100
1595	gi1518865	Sus scrofa	aminopeptidase A	876	89
1595	gi7673021	Rattus norvegicus	aminopeptidase A	847	85
1598	gi7288173	Homo sapiens	gene for alpha1,6 fucosyltransferase, exon 5.	367	100
1598	gi2055307	Homo sapiens	mRNA for N-Acetyl-beta-D-glucosaminide, complete cds.	367	100
1598	gi3451263	Homo sapiens	mRNA for glycoprotein 6-alpha-L-fucosyltransferase, transcript B1.	367	100
1599	AAB93267	Homo sapiens	Human protein sequence SEQ ID NO:12300.	149	46
1599	gi295671	Saccharomyces cerevisiae	selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependant RNA polymerase I and III	123	25
1599	gi11559313	Halocynthia roretzi	synaptotagmin	121	24
1600	AAB48139	Homo sapiens	Human TANGO 209 variant 1 polypeptide.	219	79
1600	AAB48140	Homo sapiens	Human TANGO 209 variant 2 polypeptide.	219	79
1600	AAB48141	Homo sapiens	Human TANGO 209 variant 3 polypeptide.	224	51
1602	gi12053353	Homo sapiens	mRNA; cDNA DKFZp586D0222 (from clone DKFZp586D0222); complete cds.	548	82

Table 2A  
204

SEQ ID	Hit ID	Species	Description	S score	% Identity
1602	gi7023456	Homo sapiens	cDNA FLJ11040 fis, clone PLACE1004388.	548	82
1602	gi10434241	Homo sapiens	cDNA FLJ12633 fis, clone NT2RM4001856.	548	82
1603	gi15082532	Homo sapiens	clone MGC:20434 IMAGE:4650497, mRNA, complete cds.	212	85
1603	gi13096814	Mus musculus	RIKEN cDNA 4932442K08 gene	165	72
1603	gi9622395	Plasmodium falciparum	variant surface protein	66	44
1604	gi1321596	Homo sapiens	Human (clone HSLV2132) Ig lambda light chain V region, subgroup II, partial sequence.	374	100
1604	AAG76014	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6778.	362	97
1604	AAB36212	Homo sapiens	Human immune system associated protein HISAP-10.	342	91
1605	gi10434674	Homo sapiens	cDNA FLJ12911 fis, clone NT2RP2004425, highly similar to Mus musculus axotrophin mRNA.	2340	94
1605	AAB95234	Homo sapiens	Human protein sequence SEQ ID NO:17375.	2340	94
1605	AAB27239	Homo sapiens	Human EXMAD-17 SEQ ID NO: 17.	2340	94
1606	gi8918522	Homo sapiens	PCCX2 mRNA for protein containing CXXC domain 2, partial cds.	2200	84
1606	gi14042116	Homo sapiens	cDNA FLJ14534 fis, clone NT2RM2000599, weakly similar to Homo sapiens F-box protein Lilina (LILINA) mRNA.	2446	78
1606	AAB92702	Homo sapiens	Human protein sequence SEQ ID NO:11102.	2446	78
1607	gi3176762	Homo sapiens	receptor for viral semaphorin protein (VESPR) mRNA, complete cds.	807	100
1607	AAY13462	Homo sapiens	Viral-encoded semaphorin protein receptor (VESPR) polypeptide.	807	100
1607	AAB28522	Homo sapiens	Human VESPR.	807	100
1608	gi1373425	Homo sapiens	Human bumetanide-sensitive Na-K-2Cl cotransporter (NKCC2) mRNA, complete cds.	1295	99
1608	AAW29683	Homo sapiens	Human Na-K-2Cl cotransporter NKCC2.	1295	99
1608	gi516000	Oryctolagus cuniculus	bumetanide-sensitive Na-K-Cl cotransport protein splice isoform B	1244	95
1609	AAB95655	Homo sapiens	Human protein sequence SEQ ID NO:18417.	229	61
1609	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:121.	221	58
1609	gi1196433	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	226	60
1610	AAB88388	Homo sapiens	Human membrane or secretory protein clone PSEC0131.	923	100
1610	AAB25719	Homo sapiens	Human secreted protein sequence	114	100

Table 2A  
205

SEQ ID	Hit ID	Species	Description	S score	% Identity
			encoded by gene 6 SEQ ID NO:108.		
1610	AAB25718	Homo sapiens	Human secreted protein sequence encoded by gene 6 SEQ ID NO:107.	73	100
1611	AAY04734	Homo sapiens	Protein containing PDZ domain from clone 38-2-1c.	269	98
1611	AAY04732	Homo sapiens	Protein containing PDZ domain from clone 38-2-1a.	269	98
1611	AAY53753	Homo sapiens	Amino acid sequence of the MMSC2 protein.	269	98
1612	gi1794211	Homo sapiens	Human oncostatin-M specific receptor beta subunit (OSMRB) mRNA, complete cds.	1127	92
1612	AAR85912	Homo sapiens	Oncostatin M receptor-beta subunit.	1127	92
1612	gi15012082	Homo sapiens	Similar to oncostatin M receptor, clone MGC:13583 IMAGE:4043935, mRNA, complete cds.	1127	92
1613	gi12803103	Homo sapiens	heterogeneous nuclear ribonucleoprotein A1, clone MGC:8473 IMAGE:2821751, mRNA, complete cds.	406	82
1613	gi15082486	Homo sapiens	Similar to heterogeneous nuclear ribonucleoprotein A1, clone MGC:20389 IMAGE:4564655, mRNA, complete cds.	406	82
1613	gi496898	Homo sapiens	H.sapiens mRNA for hnRNPcore protein A1.	406	82
1614	gi12653633	Homo sapiens	lysyl oxidase-like 2, clone MGC:1709 IMAGE:3347512, mRNA, complete cds.	898	99
1614	gi1890108	Homo sapiens	Human lysyl oxidase-related protein (WS9-14) mRNA, complete cds.	898	99
1614	AAB00077	Homo sapiens	Human lysyl oxidase related protein (Lor).	898	99
1615	gi12803157	Homo sapiens	COX15 (yeast) homolog, cytochrome c oxidase assembly protein, clone MGC:8634 IMAGE:2961532, mRNA, complete cds.	159	52
1615	gi15426569	Homo sapiens	clone MGC:4234 IMAGE:2961532, mRNA, complete cds.	159	52
1615	gi3603230	Homo sapiens	cytochrome oxidase assembly factor (COX15) mRNA, nuclear gene encoding mitochondrial protein, complete cds.	159	52
1616	gi1215746	Bos taurus	vacuolar system associated protein-60	1309	67
1616	gi7672979	Homo sapiens	glucosidase II beta subunit mRNA, complete cds.	1656	71
1616	gi1293640	Homo sapiens	Human protein kinase C substrate 80K-H (PRKCSH) gene, exon 15-17.	1652	70

Table 2A  
206

SEQ ID	Hit ID	Species	Description	S score	% Identity
1617	gi1418820	Homo sapiens	H.sapiens mRNA for gamma 1 isoform of 61kDa regulatory subunit of PP2A.	355	100
1617	gi1203812	Homo sapiens	Human protein phosphatase 2A B'alpha1 regulatory subunit mRNA, complete cds.	355	100
1617	gi1000892	Homo sapiens	protein phosphatase 2A B56-gamma1 (PP2A) mRNA, 3' end of cds.	355	100
1618	gi14010930	Homo sapiens	BAC clone RP11-576F1 from 2, complete sequence.	1868	100
1618	gi7022375	Homo sapiens	cDNA FLJ10379 fis, clone NT2RM2002014.	1863	99
1618	AAB92758	Homo sapiens	Human protein sequence SEQ ID NO:11220.	1863	99
1619	gi296665	Homo sapiens	Human bone marrow serine protease gene (medullasin) (leukocyte neutrophil elastase gene).	629	93
1619	gi34533	Homo sapiens	Human mRNA for medullasin (leukocyte (neutrophil) elastase.	629	93
1619	gi307123	Homo sapiens	Human elastase/medullasin mRNA, complete cds.	629	93
1620	gi9965989	Homo sapiens	calcineurin A catalytic subunit gamma isoform mRNA, complete cds.	1650	97
1620	gi258001	human, testis, mRNA, 2134 nt]. [Homo sapiens	calcineurin A catalytic subunit	1639	96
1620	gi13436077	Homo sapiens	clone MGC:10576 IMAGE:3677098, mRNA, complete cds.	1623	94
1621	gi1429374	Escherichia coli	DcuC protein	739	100
1621	gi13360118	Escherichia coli O157:H7	c4-dicarboxylate anaerobic carrier DcuC	739	100
1621	gi1786839	Escherichia coli K12	transport of dicarboxylates	739	100
1622	AAB82485	Homo sapiens	Human secretin-like receptor Zgpr1.	271	94
1622	AAB66272	Homo sapiens	Human TANGO 378 SEQ ID NO: 29.	271	94
1622	AAB82487	Homo sapiens	Human secretin-like receptor Zgpr1 splice variant.	271	94
1623	gi4062658	Escherichia coli	Flagellar basal body P-ring protein precursor	752	87
1623	gi1787320	Escherichia coli K12	homolog of Salmonella P-ring of flagella basal body	752	87
1623	gi13360922	Escherichia coli O157:H7	flagellar basal body P-ring protein FlgI	751	86
1624	gi1773192	Escherichia coli	similar to S. cerevisiae dal1	334	98
1624	gi2735238	Escherichia coli	GlxB3	334	98
1624	gi13360031	Escherichia coli O157:H7	allantoinase	333	96
1625	gi40992	Escherichia coli	dehydrogenase	1481	93
1625	gi13363752	Escherichia coli O157:H7	aspartate-semialdehyde dehydrogenase	1481	93

Table 2A  
207

SEQ ID	Hit ID	Species	Description	S score	% Identity
1625	gi3859587	Shigella sonnei	aspartate semialdehyde dehydrogenase	1481	93
1626	gi146346	Escherichia coli	phenylalanyl-tRNA synthetase beta-subunit	735	94
1626	gi1742793	Escherichia coli	Phenylalanine--tRNA ligase (EC 6.1.1.20) b chain	735	94
1626	gi13361887	Escherichia coli O157:H7	phenylalanine tRNA synthetase beta-subunit	735	94
1627	AAB49502	Homo sapiens	Clone HYASC03.	310	98
1627	gi7020468	Homo sapiens	cDNA FLJ20396 fis, clone KAT00561.	144	48
1627	AAB18980	Homo sapiens	Amino acid sequence of a human transmembrane protein.	144	48
1628	gi14021587	Mesorhizobium loti	transcriptional regulator	333	38
1628	gi14523075	Sinorhizobium meliloti	probable LysR-family protein	329	39
1628	gi9949248	Pseudomonas aeruginosa	probable transcriptional regulator	327	37
1629	gi311422	Escherichia coli	ORF-2	282	100
1629	gi15081358	uncultured bacterium	PufM	53	41
1629	gi13362968	Escherichia coli O157:H7	holin protein	52	33
1630	gi159333	Leishmania major	glycoprotein 96-92	101	28
1630	AAY91958	Homo sapiens	Human cytoskeleton associated protein 13 (CYSKP-13).	100	28
1630	gi8163686	Streptococcus pneumoniae	surface protein PspC	105	29
1632	gi887820	Escherichia coli	UUG start; possible frameshift at end?	820	71
1632	gi466651	Escherichia coli	No definition line found	343	82
1632	gi1742360	Escherichia coli	Phosphinothricin acetyltransferase (EC 2.3.1.-).	248	83
1633	gi7022678	Homo sapiens	cDNA FLJ10565 fis, clone NT2RP2002954.	290	100
1633	AAB92950	Homo sapiens	Human protein sequence SEQ ID NO:11629.	290	100
1633	gi1755198	Cavia porcellus	nitric oxide synthase	71	33
1634	AAY48563	Homo sapiens	Human breast tumour-associated protein 24.	342	100
1634	gi12804499	Homo sapiens	ribonuclease 6 precursor, clone MGC:1360 IMAGE:2959583, mRNA, complete cds.	440	66
1634	gi12804759	Homo sapiens	ribonuclease 6 precursor, clone MGC:3554 IMAGE:2959583, mRNA, complete cds.	440	66
1635	gi10438872	Homo sapiens	cDNA: FLJ22471 fis, clone HRC10529.	1620	100
1635	AAY86509	Homo sapiens	Human gene 70-encoded protein fragment, SEQ ID NO:424.	696	100
1635	AAY86510	Homo sapiens	Human gene 70-encoded protein fragment, SEQ ID NO:425.	436	100
1636	gi8096340	Homo sapiens	mRNA for RERE, complete cds.	365	97
1636	gi7413896	group A	NSP3 protein	71	29

Table 2A  
208

SEQ ID	Hit ID	Species	Description	S score	% Identity
		rotavirus			
1636	gi9368393	Human rotavirus	NSP3 protein	65	27
1637	AAB58426	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 764.	514	94
1637	AAB00191	Homo sapiens	Breast cancer protein BCN5.	514	94
1637	AAB85481	Homo sapiens	Human 23553 sulfatase polypeptide.	514	94
1638	AAAY42750	Homo sapiens	Human calcium binding protein 1 (CaBP-1).	741	100
1638	gi12060826	Homo sapiens	serologically defined breast cancer antigen NY-BR-20 mRNA, partial cds.	673	89
1638	AAAY07006	Homo sapiens	Breast cancer associated antigen precursor sequence.	637	88
1639	gi10434205	Homo sapiens	cDNA FLJ12612 fis, clone NT2RM4001582, highly similar to Mus musculus COP9 complex subunit 7b (COPS7b) mRNA.	908	86
1639	AAB94175	Homo sapiens	Human protein sequence SEQ ID NO:14484.	908	86
1639	gi15215085	Mus musculus	Similar to COP9 (constitutive photomorphogenic), subunit 7b (Arabidopsis)	900	85
1640	gi1334836	Human herpesvirus 4	BCRF2 3072 repeat, reading frame 1	98	29
1640	gi1334837	Human herpesvirus 4	BWRF1 reading frame 2	98	29
1640	gi1334838	Human herpesvirus 4	BWRF1 reading frame 3	98	29
1641	gi13161011	Homo sapiens	sclerostin gene, complete cds.	758	100
1641	gi13161020	Homo sapiens	sclerostin mRNA, complete cds.	758	100
1641	gi13236418	Homo sapiens	SOST (SOST) mRNA, complete cds.	758	100
1642	gi13436023	Mus musculus	RIKEN cDNA B230114J08 gene	279	94
1642	AAAY36125	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 510.	279	94
1642	AAAY99458	Homo sapiens	Human PRO193 amino acid sequence SEQ ID NO:410.	279	94
1643	gi15680215	Homo sapiens	Similar to ribosomal protein L23a, clone MGC:23063 IMAGE:4872948, mRNA, complete cds.	428	76
1643	gi15990426	Homo sapiens	Similar to cadherin 1, type 1, E-cadherin (epithelial), clone MGC:23017 IMAGE:3959042, mRNA, complete cds.	428	76
1643	gi404015	Homo sapiens	Human ribosomal protein L23a mRNA, partial cds.	428	76
1644	gi12653271	Homo sapiens	thyroid hormone receptor interactor 13, clone MGC:8565 IMAGE:2822981, mRNA, complete cds.	514	82
1644	gi2232019	Homo sapiens	HPV16 E1 protein binding protein mRNA, complete cds.	514	82
1644	gi703102	Homo sapiens	thyroid receptor interactor (TRIP13) mRNA, partial cds.	513	82



Table 2A

209

SEQ ID	Hit ID	Species	Description	S score	% Identity
1646	gi13544066	Homo sapiens	Similar to mini chromosome maintenance deficient 2 (S. cerevisiae), clone MGC:13220 IMAGE:3959276, mRNA, complete cds.	273	98
1646	gi14043350	Homo sapiens	clone MGC:2123 IMAGE:3143264, mRNA, complete cds.	273	98
1646	gi14044016	Homo sapiens	clone MGC:14281 IMAGE:4131943, mRNA, complete cds.	273	98
1648	gi7960207	Oncorhynchus mykiss	vitelline envelope protein alpha	252	43
1648	gi160198	Plasmodium knowlesi	circumsporozoite protein	191	43
1648	gi15384295	Mycoplasma bovis	variable surface lipoprotein Vsp422-8	177	34
1650	gi474280	Mus musculus	mannosyl-oligosaccharide alpha-1,2-mannosidase	911	88
1650	gi15929672	Mus musculus	Similar to mannosidase 1, alpha	911	88
1650	gi2154997	Sus scrofa	Man9-mannosidase	869	83
1651	gi5915662	Homo sapiens	integrin alpha 11 subunit precursor (ITGA11) mRNA, complete cds.	542	84
1651	AAB30929	Homo sapiens	Amino acid sequence of a human alpha11 integrin chain.	542	84
1651	AAB50085	Homo sapiens	Human A259.	542	84
1652	gi4512295	Homo sapiens	DNA for immunoglobulin heavy-chain variable region, complete sequence, 3 of 5.	619	100
1652	gi296657	Homo sapiens	Human Ig heavy chain gene variable region V(12G-1) (v(h)-iv family).	613	99
1652	gi185579	Homo sapiens	H.sapiens immunoglobulin germline heavy chain gene, V region.	601	97
1654	AAE04841	Homo sapiens	Human SGP039 phosphatase polypeptide.	667	92
1654	gi2582800	Medicago sativa	protein phosphatase 2C	107	41
1654	gi8778653	Arabidopsis thaliana	F9C16.6	107	40
1655	gi9928511	Mycobacterium tuberculosis	SEQ ID NO 18B'	89	27
1655	gi13897999	Galleria mellonella	silk protease inhibitor 2 precursor	50	40
1655	gi204419	Rattus norvegicus	glycam 1	85	24
1656	gi12653509	Homo sapiens	DKFZP564K1964 protein, clone MGC:8605 IMAGE:2961267, mRNA, complete cds.	227	95
1656	gi5912199	Homo sapiens	mRNA; cDNA DKFZp564K1964 (from clone DKFZp564K1964); complete cds.	227	95
1656	gi4530587	Homo sapiens	TADA1 protein mRNA, complete cds.	227	95
1658	gi12654931	Homo sapiens	protein disulfide isomerase-related protein, clone MGC:5517 IMAGE:3454007, mRNA,	2281	100

Table 2A  
210

SEQ ID	Hit ID	Species	Description	S score	% Identity
			complete cds.		
1658	gi1136743	Homo sapiens	Human mRNA for protein disulfide isomerase-related protein P5, complete cds.	2281	100
1658	AAW25154	Homo sapiens	Human disulphide epimerase like enzyme, EP52.	2281	100
1659	gi6694278	Homo sapiens	cell recognition molecule Caspr2 (CASPR2) mRNA, complete cds.	356	95
1659	gi13624214	Homo sapiens	contactin-associated protein 2 (CNTNAP2) mRNA, complete cds.	356	95
1659	gi1857708	Homo sapiens	contactin associated protein (Caspr) mRNA, complete cds.	140	37
1660	AAG02922	Homo sapiens	Human secreted protein, SEQ ID NO: 7003.	136	96
1660	gi34373	Homo sapiens	Human DNA for LINE-1 transposable element ORF1 and II.	135	96
1660	gi7959889	Homo sapiens	PRO2221	127	88
1661	gi4588087	Homo sapiens	PTH-responsive osteosarcoma B1 protein (B1) mRNA, complete cds.	569	87
1661	AAG74841	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5605.	569	87
1661	gi4588085	Homo sapiens	PTH-responsive osteosarcoma D1 protein (D1) mRNA, partial cds.	312	84
1662	gi4754907	Homo sapiens	histone deacetylase 4 mRNA, complete cds.	2090	94
1662	AAB49957	Homo sapiens	Human histone deacetylase HDAC-4.	2090	94
1662	gi14495171	Gallus gallus	histone deacetylase-4	1790	78
1663	AAB74704	Homo sapiens	Human membrane associated protein MEMAP-10.	183	100
1663	gi2621542	Methanothermobacter thermotrophicus	conserved protein	83	70
1663	gi10440349	Homo sapiens	mRNA for FLJ00009 protein, partial cds.	82	35
1664	gi15159543	Agrobacterium tumefaciens	AGR_L_2143p	99	27
1664	gi5360174	Gallus gallus	NOTCH-1	98	31
1664	gi4960212	Bos taurus	cone-rod homeobox	85	27
1665	AAB32388	Homo sapiens	Human secreted protein sequence encoded by gene 18 SEQ ID NO:74.	359	100
1665	AAV91419	Homo sapiens	Human secreted protein sequence encoded by gene 6 SEQ ID NO:140.	83	36
1665	gi860970	Homo sapiens	H.sapiens mRNA for HP8 protein.	84	35
1666	gi15012095	Homo sapiens	Similar to protease inhibitor 3, skin-derived (SKALP), clone MGC:13613 IMAGE:4083155, mRNA, complete cds.	621	100
1666	gi28712	Homo sapiens	H.sapiens encoding skin-derived antileukoproteinase.	621	100
1666	gi219615	Homo sapiens	Human gene for elafin, complete cds.	621	100

Table 2A

211

SEQ ID	Hit ID	Species	Description	S score	% Identity
1667	gi6650233	Homo sapiens	zinc finger protein 74 (ZNF74) gene, exon 3, alternative splice products and complete cds.	259	58
1667	gi15081398	Homo sapiens	kruppel-like zinc finger protein (ZNF300) mRNA, complete cds.	246	73
1667	gi1769491	Homo sapiens	Human kruppel-related zinc finger protein (ZNF184) mRNA, partial cds.	246	64
1668	AAB38280	Homo sapiens	Human secreted protein sequence encoded by gene 20 SEQ ID NO:136.	142	62
1668	gi1196431	Homo sapiens	Human factor VIII gene L1 element insertion DNA.	131	58
1668	gi6562173	Homo sapiens	mRNA; cDNA DKFZp566H033 (from clone DKFZp566H033); partial cds.	147	60
1669	AAG03136	Homo sapiens	Human secreted protein, SEQ ID NO: 7217.	179	75
1669	AAG02563	Homo sapiens	Human secreted protein, SEQ ID NO: 6644.	77	60
1669	AAG73365	Homo sapiens	Human gene 20-encoded secreted protein HFCAA91, SEQ ID NO:136.	76	63
1670	gi12804499	Homo sapiens	ribonuclease 6 precursor, clone MGC:1360 IMAGE:2959583, mRNA, complete cds.	376	83
1670	gi12804759	Homo sapiens	ribonuclease 6 precursor, clone MGC:3554 IMAGE:2959583, mRNA, complete cds.	376	83
1670	gi5091495	Homo sapiens	ribonuclease 6 precursor, mRNA, complete cds.	376	83
1672	AAG63163	Homo sapiens	Amino acid sequence of carcinoembryonic antigen-like polypeptide.	2187	99
1672	AAR54714	Homo sapiens	Carcinoembryonic antigen CEA-(c).	463	34
1672	AAB43688	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1133.	466	31
1673	gi13097624	Homo sapiens	clone IMAGE:3608084, mRNA, partial cds.	760	89
1673	gi10438279	Homo sapiens	cDNA: FLJ22029 fis, clone HEP08661.	411	51
1673	AAB93267	Homo sapiens	Human protein sequence SEQ ID NO:12300.	379	47
1674	gi339776	Homo sapiens	Human transposon L1.2.	686	85
1674	gi5070621	Homo sapiens	retrotransposon L1 insertion in X-linked retinitis pigmentosa locus, complete sequence.	686	85
1674	gi5052950	Homo sapiens	LINE1 element inserted in B-globin gene intron 2.	684	85
1675	AAB56206	Homo sapiens	Human secreted protein sequence encoded by gene 130 SEQ ID NO:300.	44	43
1675	AAB56350	Homo sapiens	Human secreted protein sequence encoded by gene 130 SEQ ID NO:444.	52	39
1675	AAW33908	Homo sapiens	Peptide analogue of human insulin-	57	31

Table 2A  
212

SEQ ID	Hit ID	Species	Description	S score	% Identity
			like growth factor-1 (IGF-1).		
1676	gi34234	Homo sapiens	H.sapiens mRNA for laminin-binding protein.	290	67
1676	gi13529269	Homo sapiens	laminin receptor 1 (67kD, ribosomal protein SA), clone MGC:12521 IMAGE:3997019, mRNA, complete cds.	289	67
1676	gi14250794	Homo sapiens	laminin receptor 1 (67kD, ribosomal protein SA), clone MGC:16750 IMAGE:4130936, mRNA, complete cds.	289	67
1677	gi475782	Homo sapiens	Human GS2 gene, exon 7 and complete cds.	1249	97
1677	gi458226	Homo sapiens	Human GS2 mRNA, complete cds.	1249	97
1677	AAG00737	Homo sapiens	Human secreted protein, SEQ ID NO: 4818.	531	94
1678	gi3089427	Homo sapiens	SSC6 rearranged T cell receptor beta chain (TCRBV17) gene, complete cds.	444	100
1678	gi3089433	Homo sapiens	SSC9 rearranged T cell receptor beta chain (TCRBV17) gene, complete cds.	444	100
1678	gi3002927	Homo sapiens	T cell receptor beta chain (TCRBV17S1-TCRBJ1S5) mRNA, complete cds.	444	100
1679	gi15929119	Homo sapiens	clone MGC:8834 IMAGE:3920437, mRNA, complete cds.	455	100
1679	gi7022159	Homo sapiens	cDNA FLJ10242 fis, clone HEMBB1000630.	455	100
1679	AAB92624	Homo sapiens	Human protein sequence SEQ ID NO:10919.	455	100
1680	gi10186503	Homo sapiens	sialic acid-specific acetyltransferase II mRNA, complete cds, alternatively spliced.	817	100
1680	gi6808138	Homo sapiens	mRNA; cDNA DKFZp761A051 (from clone DKFZp761A051); partial cds.	817	100
1680	gi10242345	Homo sapiens	sialic acid-specific 9-O-acetyltransferase I mRNA, complete cds.	753	100
1681	gi4768463	Homo sapiens	clone 118 T-cell receptor beta chain (TCRBV10S1P) mRNA, partial cds.	76	44
1681	gi4768574	Homo sapiens	clone 179 T-cell receptor beta chain (TCRBV10S1P) mRNA, partial cds.	71	42
1681	gi4768459	Homo sapiens	clone 115 T-cell receptor beta chain (TCRBV10S1P) mRNA, partial cds.	69	44
1682	gi4097459	Elephantulus edwardii	reverse transcriptase	68	48
1683	gi1220315	Homo sapiens	early placenta insulin-like peptide EPIL (INSL4) mRNA, complete cds.	349	100
1683	AAR89134	Homo sapiens	Human early placental insulin-like protein.	349	100

Table 2A  
213

SEQ ID	Hit ID	Species	Description	S score	% Identity
1683	AAW17676	Homo sapiens	Human relaxin-related factor-2 (RRF-2).	349	100
1684	gi23365	Homo sapiens	Human mRNA for 17-beta-hydroxysteroid dehydrogenase (17-HSD) (EC 1.1.1.62).	800	83
1684	gi975895	Homo sapiens	Human 17-beta-hydroxysteroid dehydrogenase (EDH17B2) gene, complete cds.	800	83
1684	gi177127	Homo sapiens	Human 17-beta-hydroxysteroid dehydrogenase (EDH17B1 and EDH17B2) genes, complete coding regions and flanks.	800	83
1685	AAB18919	Homo sapiens	A novel polypeptide designated PRO4356.	1336	99
1685	AAB31206	Homo sapiens	Amino acid sequence of human polypeptide PRO4356.	1336	99
1685	gi5834584	Homo sapiens	mRNA encoding rat C4.4-like protein.	231	31
1686	gi1655963	Homo sapiens	Human transforming growth factor-beta type II receptor gene, exon 7 and complete cds.	1634	99
1686	gi339570	Homo sapiens	Human TGF-beta type II receptor mRNA, complete cds.	1634	99
1686	AAR36601	Homo sapiens	TGF-beta1 receptor type II (clone 3FF).	1634	99
1687	gi508260	Homo sapiens	Human type I vasoactive intestinal peptide receptor (V1RG) gene, exon 13 and complete cds.	390	76
1687	gi407462	Homo sapiens	H.sapiens HIVR mRNA for vasoactive intestinal peptide (VIP) receptor.	390	76
1687	gi292904	Homo sapiens	Human vasoactive intestinal peptide receptor mRNA, complete cds.	390	76
1688	gi181268	Homo sapiens	Human c-yes-1 mRNA.	520	90
1688	AAY24421	Homo sapiens	Human yes1 protein.	520	90
1688	AAB84663	Homo sapiens	Amino acid sequence of human tyrosine kinase protein Yes.	520	90
1689	AAG02314	Homo sapiens	Human secreted protein, SEQ ID NO: 6395.	139	55
1689	gi13325174	Homo sapiens	clone MGC:10997 IMAGE:3638158, mRNA, complete cds.	104	70
1689	gi9652123	Mus musculus	disrupter of silencing SAS10	86	40
1690	gi13516467	Homo sapiens	HDNB1(homzygously deleted in neuroblastoma-1)/UFD2 mRNA, complete cds.	1217	89
1690	gi4104976	Homo sapiens	ubiquitin-fusion degradation protein 2 (UFD2) mRNA, complete cds.	1217	89
1690	gi14582754	Homo sapiens	ubiquitin-fusion degradation protein 2 mRNA, complete cds.	1217	89
1691	gi2463632	Homo sapiens	monocarboxylate transporter homologue MCT6 mRNA, complete cds.	326	96
1691	AAG73420	Homo sapiens	Human gene 21-encoded secreted protein HBWBR94, SEQ ID NO:192.	63	35

Table 2A

214

SEQ ID	Hit ID	Species	Description	S score	% Identity
1691	gi9246437	Staphylococcus aureus	fntA-like protein	83	26
1692	AAB58175	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 513.	394	98
1692	gi14211500	Homo sapiens	secretory protein SEC8 mRNA, complete cds.	394	98
1692	gi14042555	Homo sapiens	cDNA FLJ14782 fis, clone NT2RP4000524, highly similar to Mus musculus Sec8 mRNA.	394	98
1693	gi3786312	Homo sapiens	mRNA for extracellular matrix protein, complete cds.	498	39
1693	gi13937865	Homo sapiens	lumican, clone MGC:12410 IMAGE:3950745, mRNA, complete cds.	312	29
1693	gi699577	Homo sapiens	Human lumican mRNA, complete cds.	312	29
1694	gi3002588	Mus musculus	Plenty of SH3s; POSH	219	93
1694	gi7230620	Rattus norvegicus	SH3 domain-containing adapter protein isoform SETA-1x23	130	34
1694	AAW34246	Homo sapiens	SH3 domain of human clone 53 protein.	114	50
1695	gi10645308	Leishmania major	L8453.1	151	27
1695	gi15419013	Toxoplasma gondii	subtilisin-like protein	147	30
1695	gi12018147	Chlamydomonas reinhardtii	vegetative cell wall protein gp1	143	30
1696	AAB43791	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1236.	353	98
1697	gi156368	Caenorhabditis elegans	metallothionein-2	56	39
1697	gi156381	Caenorhabditis elegans	metallothionein-like protein	56	39
1697	gi6782	Caenorhabditis elegans	metallothionein-II	56	39
1698	gi9858855	Homo sapiens	HPT protein (HPT) mRNA, complete cds; nuclear gene for mitochondrial product.	1318	85
1698	AAB29653	Homo sapiens	Human membrane-associated protein HUMAP-10.	1318	85
1698	AAB32389	Homo sapiens	Human secreted protein sequence encoded by gene 19 SEQ ID NO:75.	1318	85
1699	gi6841138	Homo sapiens	HSPC099 mRNA, partial cds.	275	100
1699	gi7022824	Homo sapiens	cDNA FLJ10656 fis, clone NT2RP2006038.	123	23
1699	AAB93037	Homo sapiens	Human protein sequence SEQ ID NO:11816.	123	23
1700	AAB36587	Homo sapiens	Human FLEXHT-9 protein sequence SEQ ID NO:9.	585	53
1700	gi7023841	Homo sapiens	cDNA FLJ11280 fis, clone PLACE1009459.	581	54
1700	AAB93652	Homo sapiens	Human protein sequence SEQ ID NO:13161.	581	54
1702	gi12330618	Human immunodeficiency	envelope glycoprotein	71	29

Table 2A  
215

SEQ ID	Hit ID	Species	Description	S score	% Identity
		y virus type 1			
1703	AAR48975	Homo sapiens	Human beta1,6-N-acetylglucosaminyltransferase protein.	413	100
1703	gi870752	Homo sapiens	Human mRNA for N-acetylglucosaminyltransferase V, complete cds.	413	100
1703	gi4545222	Homo sapiens	alpha-1,3(6)-mannosylglycoprotein beta-1,6-N-acetylglucosaminyltransferase (MGAT5) mRNA, complete cds.	413	100
1704	gi12654535	Homo sapiens	HSPC025, clone MGC:735 IMAGE:3507964, mRNA, complete cds.	1251	81
1704	gi13960140	Homo sapiens	HSPC025, clone MGC:4223 IMAGE:2959747, mRNA, complete cds.	1251	81
1704	gi4679028	Homo sapiens	HSPC021	1251	81
1706	gi14250636	Homo sapiens	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 2, clone MGC:3398 IMAGE:3628374, mRNA, complete cds.	2605	86
1706	gi6580428	Homo sapiens	partial NFKBIL2 gene for IkappaBR, exons 1-13.	2751	95
1706	gi746415	Homo sapiens	Human I kappa BR mRNA, complete cds.	1668	72
1707	AAB95830	Homo sapiens	Human protein sequence SEQ ID NO:18850.	219	72
1707	gi7959889	Homo sapiens	PRO2221	137	49
1707	gi2072969	Homo sapiens	Human L1 element L1.24 p40 gene, complete cds.	133	48
1708	gi5901529	Homo sapiens	C2H2 type Kruppel-like zinc finger protein splice variant b (ZNF236) mRNA, complete cds.	565	99
1708	gi5705917	human, MOLT 4 T-cells, mRNA, 798 nt]. [Homo sapiens	HKR-T1=Kruppel-like zinc finger protein	223	45
1708	gi498736	Homo sapiens	H.sapiens HZF9 mRNA for zinc finger protein.	220	48
1709	gi1865716	Bos taurus	procollagen I N-proteinase	265	42
1709	AAW47029	Homo sapiens	Human N-proteinase (70 kDa short form).	254	43
1709	AAW47030	Homo sapiens	Bovine N-proteinase.	254	42
1710	gi12862392	Mus musculus	D86	1379	78
1710	gi763113	Homo sapiens	H.sapiens ERK3 mRNA.	872	92
1710	gi1294779	Homo sapiens	ERK3 protein kinase mRNA, complete cds.	872	92
1711	gi572680	Escherichia coli	beta ketoacyl-acyl carrier protein synthase	315	91
1711	gi664870	Escherichia coli	beta-ketoacyl-acyl carrier protein synthase II	315	91
1711	gi4062664	Escherichia coli	3-oxoacyl- synthase (EC 2.3.1.41) II	315	91
1712	AAR89952	Homo sapiens	Insulin-like growth factor binding protein-3.	726	95

Table 2A

216

SEQ ID	Hit ID	Species	Description	S score	% Identity
1712	gi12652547	Homo sapiens	insulin-like growth factor binding protein 3, clone MGC:2305 IMAGE:3506666, mRNA, complete cds.	722	95
1712	gi398164	Homo sapiens	H.sapiens mRNA for insulin-like growth factor binding protein-3.	722	95
1713	gi10435700	Homo sapiens	cDNA FLJ13633 fis, clone PLACE1011114, weakly similar to PROBABLE ATP-DEPENDENT RNA HELICASE HAS1.	710	97
1713	AAB99891	Homo sapiens	Human RNA helicase gene helicain B protein sequence SEQ ID NO:4.	710	97
1713	AAB99892	Homo sapiens	Human RNA helicase gene helicain C protein sequence SEQ ID NO:6.	710	97
1714	gi15384740	Homo sapiens	mRNA for Paralemmin-2 (PALM2 gene).	1652	100
1714	gi15384742	Homo sapiens	mRNA for Palm2-AKAP2 fusion protein (PALM2/AKAP2 gene).	1577	100
1714	gi14041780	Homo sapiens	mRNA for AKAP-2 protein.	415	60
1715	gi3600073	Homo sapiens	Human endogenous retrovirus K clone 7.1 polymerase mRNA, partial cds.	573	48
1715	gi3600067	Homo sapiens	Human endogenous retrovirus K clone 10.9 polymerase mRNA, partial cds.	572	48
1715	gi1780973	Human endogenous retrovirus K	pol protein	572	48
1716	AAU00025	Homo sapiens	Human activated T-lymphocyte associated sequence 4, ATLAS-4.	315	74
1716	gi5880909	Drosophila melanogaster	RNA-binding protein Smaug	87	29
1716	gi7380929	Drosophila melanogaster	smaug protein	87	29
1717	gi35825	Homo sapiens	Human mRNA for pregnancy zone protein.	3998	93
1717	gi579594	Homo sapiens	alpha 2-macroglobulin 690-740	2841	64
1717	AAR11749	Homo sapiens	Human alpha-2 macroglobulin bait region mutant.	2832	64
1718	gi4760578	Mus musculus	PB-Cadherin	1033	84
1718	gi1398906	Rattus norvegicus	long type PB-cadherin	1027	84
1718	gi1398912	Rattus norvegicus	short type PB-cadherin	1027	84
1719	gi339776	Homo sapiens	Human transposon L1.2.	312	62
1719	gi5070621	Homo sapiens	retrotransposon L1 insertion in X-linked retinitis pigmentosa locus, complete sequence.	312	62
1719	gi7959889	Homo sapiens	PRO2221	308	71
1720	AAE06588	Homo sapiens	Human protein having hydrophobic domain, HP10778.	687	100
1720	gi15072402	Raja erinacea	organic solute transporter alpha	357	45
1720	AAB38348	Homo sapiens	Human secreted protein encoded by gene 28 clone HLDOW79.	328	100



Table 2A

217

SEQ ID	Hit ID	Species	Description	S score	% Identity
1721	AAG81345	Homo sapiens	Human AFP protein sequence SEQ ID NO:208.	525	90
1721	AAB93797	Homo sapiens	Human protein sequence SEQ ID NO:13560.	525	90
1721	AAB44681	Homo sapiens	Human secreted protein sequence encoded by gene 41 SEQ ID NO:146.	199	100
1722	gi1809225	Homo sapiens	Human BAC clone RG161K23 from 7q21, complete sequence.	2183	83
1722	gi1698396	Homo sapiens	Human lanosterol 14-demethylase cytochrome P450 (CYP51) mRNA, complete cds.	2183	83
1722	gi871883	Homo sapiens	Human mRNA for lanosterol 14-demethylase, complete cds.	2183	83
1723	gi16182326	Drosophila melanogaster	GH01206p	108	29
1723	AAG77172	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:7938.	90	39
1723	gi3878064	Caenorhabditis elegans	H19J13.1	72	25
1724	AAY53040	Homo sapiens	Human secreted protein clone kj320_1 protein sequence SEQ ID NO:86.	2480	100
1724	gi3510639	Rattus norvegicus	UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase T5	1351	59
1724	gi6688167	Homo sapiens	partial mRNA for GalNAc-T5 (GALNT5 gene).	1082	100
1725	gi14603092	Homo sapiens	Similar to CD47 antigen (Rh-related antigen, integrin-associated signal transducer), clone MGC:15298 IMAGE:4303534, mRNA, complete cds.	1329	97
1725	gi15277580	Homo sapiens	clone MGC:9240 IMAGE:3857911, mRNA, complete cds.	1329	97
1725	gi396705	Homo sapiens	H.sapiens integrin associated protein mRNA, complete CDS.	1329	97
1726	gi1864011	Homo sapiens	mRNA for SHPS-1, complete cds.	858	98
1726	gi2052056	Homo sapiens	H.sapiens mRNA for SIRP-alpha1.	858	98
1726	gi6518913	Homo sapiens	Bit mRNA, complete cds.	858	98
1727	gi2707601	Homo sapiens	A4 differentiation-dependent protein (A4), triple LIM domain protein (LMO6), and synaptophysin (SYP) genes, complete cds; and calcium channel alpha-1 subunit (CACNA1F) gene, partial cds.	1656	100
1727	gi6180176	Homo sapiens	transcription factor IGHM enhancer 3, JM11 protein, JM4 protein, JM5 protein, T54 protein, JM10 protein, A4 differentiation-dependent protein, triple LIM domain protein 6, and synaptophysin genes, complete cds; and L-type calcium channel alpha-1 subunit gene, partial cds, complete sequence.	1656	100

Table 2A  
218

SEQ ID	Hit ID	Species	Description	S score	% Identity
1727	gi899301	Homo sapiens	Human mRNA for synaptophysin (p38).	1602	100
1728	gi15590682	Homo sapiens	histone deacetylase 9a (HDAC9) mRNA, complete cds, alternatively spliced.	629	96
1728	gi12060992	Mus musculus	MEF2-interacting transcription repressor MITR	626	95
1728	gi13183017	Mus musculus	histone deacetylase-related protein	623	94
1729	gi5911884	Homo sapiens	mRNA; cDNA DKFZp434N126 (from clone DKFZp434N126).	1605	99
1729	gi15912209	Arabidopsis thaliana	Atlg27520/T17H3_2	228	34
1729	gi14164377	Mus musculus	Type II membrane protein of ER-mouse gene similar to alpha-mannosidase	216	38
1730	gi21842	Triticum aestivum	proline-rich protein	403	33
1730	gi4138732	Zea mays	proline-rich protein	398	30
1730	gi11610622	Rattus norvegicus	GABA-A epsilon subunit splice variant	411	32
1731	gi189222	Homo sapiens	Human neurokinin-2 receptor (TAC2R) gene, exon 5.	1733	95
1731	AAW80135	Homo sapiens	Human recombinant neurokinin-2 (NK-2) receptor protein.	1733	95
1731	gi189135	Homo sapiens	Human neurokinin A receptor (NK-2R) mRNA, complete cds.	1732	95
1732	AAB75594	Homo sapiens	Human secreted protein sequence encoded by gene 37 SEQ ID NO:148.	678	99
1732	AAB75542	Homo sapiens	Human secreted protein sequence encoded by gene 37 SEQ ID NO:96.	294	100
1732	gi1864011	Homo sapiens	mRNA for SHPS-1, complete cds.	261	43
1733	AAY66648	Homo sapiens	Membrane-bound protein PRO1120.	294	93
1733	AAB65171	Homo sapiens	Human PRO1120 (UNQ559) protein sequence SEQ ID NO:84.	294	93
1733	AAE01440	Homo sapiens	Human gene 5 encoded secreted protein HE9QN39, SEQ ID NO:95.	294	93
1734	gi13161409	Mus musculus	family 4 cytochrome P450	539	78
1734	gi155947	Blaberus discoidalis	cytochrome P450	248	43
1734	gi3249041	Diploptera punctata	corpora allata cytochrome P450	233	42
1735	AAY71159	Homo sapiens	Human phosphodiesterase interacting protein, myomegalin.	682	45
1735	gi4761644	Rattus norvegicus	myomegalin	484	55
1735	gi53782	Mus musculus	profilin (AA 1-140)	414	74
1736	gi4959382	Homo sapiens	human endogenous retrovirus HERV-H19 pol protein (pol) gene, partial cds; env protein (env) gene, complete cds; and 3' LTR, complete sequence.	301	39
1736	gi8439396	HERV-H/env62	envelope protein	301	39

Table 2A  
219

SEQ ID	Hit ID	Species	Description	S score	% Identity
1736	gi5442112	Simian retrovirus type 2	envelope glycoprotein	296	42
1737	AAY76177	Homo sapiens	Human secreted protein encoded by gene 54.	288	100
1738	gi3688630	Homo sapiens	hdclk-4 gene, exon3, exon4 and complete cds.	759	100
1738	gi6049610	Homo sapiens	dickkopf-4 (DKK-4) mRNA, complete cds.	759	100
1738	AAW73017	Homo sapiens	Human cysteine-rich secreted protein CRSP-2.	759	100
1739	gi10439926	Homo sapiens	cDNA: FLJ23293 fis, clone HEP10514.	1786	99
1739	gi4406632	Homo sapiens	clone 25221 mRNA sequence, complete cds.	1253	64
1739	gi10435296	Homo sapiens	cDNA FLJ13321 fis, clone OVARC1001703, weakly similar to Mus musculus ARL-6 interacting protein-2 (Aip-2) mRNA.	1084	62
1740	AAB49278	Homo sapiens	Protein encoded by zsig81 cDNA fragment.	755	94
1740	gi1335033	Homo sapiens	Human mRNA for collagen VI alpha-2 alternative C-terminal domain.	69	24
1740	AAY40063	Homo sapiens	Peptide sequence derived from a human secreted protein.	63	35
1741	gi14714807	Mus musculus	Similar to transporter-like protein	565	83
1741	AAG81264	Homo sapiens	Human AFP protein sequence SEQ ID NO:46.	657	100
1741	AAY66673	Homo sapiens	Membrane-bound protein PRO1115.	657	100
1742	gi14583077	Homo sapiens	PAS-kinase (PASK) mRNA, complete cds.	2227	99
1742	AAB65630	Homo sapiens	Novel protein kinase, SEQ ID NO: 156.	2223	99
1742	gi15487238	Mus musculus	PASKIN protein	1632	72
1743	gi13879899	Mycobacterium tuberculosis CDC1551	PPE family protein	118	31
1743	gi1334643	Xenopus laevis	APEG precursor protein	90	34
1743	gi184511	Homo sapiens	Human zinc-finger DNA-binding motifs (IA-1) mRNA, complete cds.	86	32
1744	gi5931718	Chlamydomonas reinhardtii	1-alpha dynein heavy chain	1874	62
1744	gi9409781	Chlamydomonas reinhardtii	1 beta dynein heavy chain	986	36
1744	gi514215	Chlamydomonas reinhardtii	dynein beta heavy chain	978	36
1745	AAY13392	Homo sapiens	Amino acid sequence of protein PRO328.	1168	57
1745	AAB01373	Homo sapiens	Neuron-associated protein.	1168	57
1745	AAY95343	Homo sapiens	Human PRO328 antitumour protein.	1168	57
1746	AAY99398	Homo sapiens	Human PRO1301 (UNQ667) amino acid sequence SEQ ID	677	94

Table 2A  
220

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NO:212.		
1746	AAW88501	Homo sapiens	Human stomach carcinoma clone HP10415-encoded protein.	675	94
1746	AAB24255	Homo sapiens	Human cytochrome P450 (HUCYP) protein SEQ ID NO:1.	675	94
1747	AAG02314	Homo sapiens	Human secreted protein, SEQ ID NO: 6395.	78	66
1747	AAB45051	Homo sapiens	Human secreted protein encoded by gene 27.	60	30
1747	gi2585662	Human immunodeficiency virus type 1	envelope glycoprotein	58	38
1748	AAE03560	Homo sapiens	Human differentially expressed kidney cDNA 22360 encoded protein.	1142	88
1748	gi15637151	Beta vulgaris	glycine decarboxylase subunit P	62	36
1749	gi13506805	Homo sapiens	thymic stromal co-transporter mRNA, complete cds.	2309	100
1749	gi13506808	Mus musculus	thymic stromal co-transporter	1789	77
1749	AAE04906	Homo sapiens	Human transporter and ion channel-19 (TRICH-19) protein.	305	34
1750	gi10438815	Homo sapiens	cDNA: FLJ22427 fis, clone HRC09013.	4170	98
1750	AAB01381	Homo sapiens	Neuron-associated protein.	2964	95
1750	gi6650377	Mus musculus	pecanex 1	2288	73
1751	gi825663	Homo sapiens	H.sapiens GLAST1 gene for glial glutamate transporter, exon1, exon 2.	411	100
1751	gi487339	Homo sapiens	Human excitatory amino acid transporter1 mRNA, complete cds.	411	100
1751	gi825504	Homo sapiens	Human mRNA for glutamate transporter, complete cds.	411	100
1752	gi1621607	Homo sapiens	Human neogenin mRNA, complete cds.	593	100
1752	gi641966	Gallus gallus	neogenin	591	98
1752	gi1785999	Rattus norvegicus	neogenin	586	97
1753	gi7020927	Homo sapiens	cDNA FLJ20674 fis, clone KAIA4450.	293	25
1753	AAP94014	Homo sapiens	Carcinoembryonic cell surface antigen.	254	27
1753	AAR60619	Homo sapiens	Carcinoembryonic antigen glycoprotein.	250	27
1754	AAY19507	Homo sapiens	Amino acid sequence of a human secreted protein.	343	91
1754	AAY19654	Homo sapiens	SEQ ID NO 372 from WO9922243.	85	100
1754	gi15075730	Sinorhizobium meliloti	HYPOTHETICAL TRANSMEMBRANE PROTEIN	65	30
1755	gi2506078	Mus musculus	tetracycline transporter-like protein	1120	99
1755	AAY29332	Homo sapiens	Human secreted protein clone pe584_2 protein sequence.	806	78
1755	AAB75295	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:114.	806	78
1756	AAE06608	Homo sapiens	Human protein having	1065	98

Table 2A

221

SEQ ID	Hit ID	Species	Description	S score	% Identity
			hydrophobic domain, HP10798.		
1756	AAB88469	Homo sapiens	Human membrane or secretory protein clone PSEC0027.	1065	98
1756	gi9664030	Drosophila melanogaster	aquaporin	115	22
1757	gi8925284	Homo sapiens	phosphatidylinositol polyphosphate 5-phosphatase type IV mRNA, complete cds.	2598	99
1757	gi9295353	Mus musculus	inositol polyphosphate 5-phosphatase	1993	74
1757	gi5360761	Rattus norvegicus	pharbin	1933	77
1758	gi395207	Bos taurus	potassium channel (BGK5)	757	83
1758	gi186669	Homo sapiens	Human potassium channel mRNA, complete cds.	754	83
1758	gi304652	Canis familiaris	delayed rectifier K <sup>+</sup> channel	750	82
1759	gi7023003	Homo sapiens	cDNA FLJ10769 fis, clone NT2RP4000151.	647	85
1759	AAB93147	Homo sapiens	Human protein sequence SEQ ID NO:12057.	647	85
1759	AAB38451	Homo sapiens	Fragment of human secreted protein encoded by gene 23 clone HEGAK44.	484	84
1760	AAB87763	Homo sapiens	Human T2R33 amino acid sequence SEQ ID NO:56.	693	83
1760	AAB87780	Homo sapiens	Human T2R50 amino acid sequence SEQ ID NO:76.	671	82
1760	AAE03828	Homo sapiens	Human gene 11 encoded secreted protein HHAUQ28, SEQ ID NO: 74.	656	89
1761	gi13543624	Homo sapiens	uncharacterized hematopoietic stem/progenitor cells protein MDS029, clone MGC:14612 IMAGE:4051044, mRNA, complete cds.	514	87
1761	gi13937872	Homo sapiens	uncharacterized hematopoietic stem/progenitor cells protein MDS029, clone MGC:12437 IMAGE:3930701, mRNA, complete cds.	514	87
1761	gi14250122	Homo sapiens	uncharacterized hematopoietic stem/progenitor cells protein MDS029, clone MGC:14755 IMAGE:4283253, mRNA, complete cds.	514	87
1762	gi2739094	Homo sapiens	sodium/myo-inositol cotransporter (SLC5A3) gene, complete cds.	3621	100
1762	gi662843	Homo sapiens	Na <sup>+</sup> /myo-inositol cotransporter (SLC5A3) gene, complete cds.	3616	99
1762	gi1237437	Bos taurus	Na <sup>+</sup> /myo-inositol cotransporter	3486	95
1763	AAW84596	Homo sapiens	Amino acid sequence of the human Tango-79 protein.	1212	44
1763	gi15029689	Homo sapiens	clone MGC:17422 IMAGE:4214343, mRNA, complete cds.	1209	44
1763	AAB74705	Homo sapiens	Human membrane associated protein MEMAP-11.	1209	44

Table 2A  
222

SEQ ID	Hit ID	Species	Description	S score	% Identity
1764	gi5923786	Homo sapiens	zinc metalloprotease ADAMTS6 (ADAMTS6) mRNA, complete cds.	614	43
1764	AAB72282	Homo sapiens	Human ADAMTS-6 amino acid sequence.	614	43
1764	gi12053709	Homo sapiens	mRNA for ADAMTS12.	1078	41
1765	gi1695682	Homo sapiens	mRNA for hepatic triglyceride lipase, complete cds.	152	64
1765	gi32498	Homo sapiens	Human mRNA for hepatic triglyceride lipase (HTGL).	152	64
1765	gi307129	Homo sapiens	Human hepatic lipase mRNA, complete cds.	152	64
1766	gi886282	Homo sapiens	glycoprotein Ib alpha (GPIb) gene, partial cds.	48	42
1766	AAB64868	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:154.	66	39
1766	gi10798865	Homo sapiens	zinc finger transcription factor BTEB2 gene, partial cds.	65	29
1767	gi8886005	Homo sapiens	lysophosphatidic acid acyltransferase-delta (LPAAT-delta) mRNA, complete cds.	324	90
1767	AAY96592	Homo sapiens	Human lysophosphatidic acid acyltransferase delta.	324	90
1767	AAY66665	Homo sapiens	Membrane-bound protein PRO1016.	324	90
1768	gi1160183	Homo sapiens	H.sapiens mRNA for metabotropic glutamate receptor type 4.	926	99
1768	gi1935039	Homo sapiens	Human metabotropic glutamate receptor 4 mRNA, complete cds.	926	99
1768	gi2298840	unidentified	HMGLUR4	926	99
1769	gi13279140	Homo sapiens	Similar to synaptotagmin 11, clone MGC:10881 IMAGE:3621175, mRNA, complete cds.	2210	99
1769	gi15489165	Homo sapiens	clone MGC:17226 IMAGE:4179089, mRNA, complete cds.	2210	99
1769	gi14042290	Homo sapiens	cDNA FLJ14634 fis, clone NT2RP2001081, moderately similar to SYNAPTOTAGMIN IV.	2201	98
1770	gi14249942	Homo sapiens	Similar to RIKEN cDNA 0610008P16 gene, clone MGC:15937 IMAGE:3537224, mRNA, complete cds.	239	78
1770	AAB56487	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1065.	239	78
1770	AAB73512	Homo sapiens	Human transferase HTFS-19, SEQ ID NO:19.	239	78
1771	gi7678873	Homo sapiens	mRNA for vascular cadherin-2, complete cds.	5369	99
1771	gi7407150	Homo sapiens	protocadherin 12 (PCDH12) mRNA, complete cds.	5369	99
1771	gi8164037	Homo sapiens	vascular endothelial cadherin 2 mRNA, complete cds.	5369	99

Table 2A  
223

SEQ ID	Hit ID	Species	Description	S score	% Identity
1772	gi15082281	Homo sapiens	Similar to steroid dehydrogenase homolog, clone MGC:10252 IMAGE:3622879, mRNA, complete cds.	231	93
1772	gi15214803	Homo sapiens	Similar to steroid dehydrogenase homolog, clone MGC:13329 IMAGE:4281565, mRNA, complete cds.	231	93
1772	gi5531815	Homo sapiens	steroid dehydrogenase homolog	231	93
1773	AAG73650	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:4414.	265	58
1773	AAB94891	Homo sapiens	Human protein sequence SEQ ID NO:16231.	265	64
1773	gi1335205	Homo sapiens	Human DNA for LINE-1 transposable element ORFI and II.	273	48
1774	AAW71708	Homo sapiens	Human integral membrane protein TMP-1.	657	50
1774	gi3603459	Homo sapiens	tetraspan NET-5 mRNA, complete cds.	626	90
1774	AAB93885	Homo sapiens	Human protein sequence SEQ ID NO:13815.	626	90
1775	AAG71597	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1278.	981	99
1775	AAG71587	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1268.	746	75
1775	gi15293749	Homo sapiens	clone OR10R2 olfactory receptor gene, partial cds.	725	75
1776	AAB38011	Homo sapiens	Human secreted protein encoded by gene 3 clone HPJXC13.	252	85
1776	AAB64888	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:66.	252	85
1776	gi10432803	Homo sapiens	cDNA FLJ11531 fis, clone HEMBA1002661.	192	54
1777	gi182851	Homo sapiens	Human G0S2 protein gene, complete cds.	497	100
1777	gi182853	Homo sapiens	Human GOS2 gene, 5' flank and cds.	497	100
1777	gi1213013	Mus musculus	G0S2-like protein	377	77
1778	gi4027903	Homo sapiens	VAMP5 mRNA, complete cds.	473	100
1778	gi4679008	Homo sapiens	VAMP5-like protein	473	100
1778	AAW04181	Homo sapiens	Cellubrevin-2.	473	100
1779	gi5326919	Bos taurus	SCO-spondin	75	30
1779	gi3059229	Aspergillus oryzae	HAPB	74	35
1779	AAAY84596	Homo sapiens	Fragment of human pre-pro-artemin polypeptide.	48	40
1780	gi1399321	Macaca mulatta	MHC class I antigen Mamu B*08	679	72
1780	gi1399319	Macaca mulatta	MHC class I antigen Mamu B*07	674	67
1780	gi8117799	Pan troglodytes	MHC class I antigen	677	70
1781	gi15929602	Homo sapiens	clone MGC:17861 IMAGE:3903313, mRNA, complete cds.	607	51
1781	gi15099953	Homo sapiens	diacylglycerol acyltransferase 2 mRNA, complete cds.	607	51
1781	AAAY99421	Homo sapiens	Human PRO1433 (UNQ738) amino acid sequence SEQ ID	607	51

Table 2A  
224

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NO:292.		
1782	gi2706518	Homo sapiens	PMP69 gene, exon 1 (and joined CDS).	1602	98
1782	gi2343157	Homo sapiens	peroxisomal membrane protein 69 (PMP69) mRNA, complete cds.	1602	98
1782	gi15215442	Homo sapiens	Similar to ATP-binding cassette, sub-family D (ALD), member 4, clone MGC:4125 IMAGE:2960427, mRNA, complete cds.	1594	98
1783	AAB35235	Homo sapiens	Human neurotransmitter transporter protein GC42.	3571	98
1783	gi546769	human, substantia nigra, mRNA, 2364 nt]. [Homo sapiens	glycine transporter type 1b	3557	98
1783	AAB35236	Homo sapiens	Human glycine transporter type 1c.	3557	98
1784	AAU00017	Homo sapiens	Human Plexin-D1.	3361	97
1784	gi13097621	Homo sapiens	clone IMAGE:3607457, mRNA, partial cds.	2470	100
1784	AAU00015	Homo sapiens	Human Plexin-B2.	1290	47
1785	gi619726	Homo sapiens	Human nuclear factor I (NFI) mRNA, clone AT1, complete cds.	953	99
1785	gi619730	Homo sapiens	Human nuclear factor I (NFI) mRNA, clone CT1, partial cds.	953	99
1785	gi305357	Mesocricetus auratus	nuclear factor 1-like protein	953	99
1786	gi13810568	Homo sapiens	TLR5 mRNA for Toll-like receptor 5, complete cds.	4482	100
1786	gi3132526	Homo sapiens	Toll/interleukin-1 receptor-like protein 3 (TIL3) mRNA, complete cds.	4464	99
1786	gi7648687	Mus musculus	toll-like receptor 5	3235	72
1787	AAB56473	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1051.	474	100
1787	AAB60119	Homo sapiens	Human transport protein TPPT-39.	456	98
1787	gi11463949	Homo sapiens	hUGTrel7 mRNA for UDP-glucuronic acid, complete cds.	308	62
1788	gi4929765	Homo sapiens	CGI-148 protein mRNA, complete cds.	744	89
1788	gi7578785	Homo sapiens	NPD008 protein (NPD008) mRNA, complete cds.	744	89
1788	gi14250060	Homo sapiens	clone MGC:14598 IMAGE:4292664, mRNA, complete cds.	737	88
1789	gi1736785	Escherichia coli	Acriflavin resistance protein F (EnvD protein).	2265	100
1789	gi15980819	Yersinia pestis	AcrB/AcrD/AcrF family membrane protein	1854	79
1789	gi1736782	Escherichia coli	Acriflavin resistance protein F (EnvD protein).	1821	100
1790	gi15160166	Agrobacterium tumefaciens	AGR_L_3181p	549	59
1790	gi4981492	Thermotoga maritima	ribose ABC transporter, permease protein	316	40



Table 2A  
225

SEQ ID	Hit ID	Species	Description	S score	% Identity
1790	gi12724646	Lactococcus lactis subsp. lactis	ribose ABC transporter permease protein	314	35
1791	gi41587	Escherichia coli	glycerol-3-phosphatase transporter (AA 1 - 452, glpT)	1073	100
1791	gi1799587	Escherichia coli	glycerol-3-phosphate transport protein	1073	100
1791	gi1788573	Escherichia coli K12	sn-glycerol-3-phosphate permease	1073	100
1792	gi148200	Escherichia coli	similar to arylsulfatases of Klebsiella pneumoniae and Homo sapiens	679	98
1792	gi13364207	Escherichia coli O157:H7	arylsulfatase	679	98
1792	gi1790233	Escherichia coli K12	arylsulfatase	679	98
1793	gi9657461	Vibrio cholerae	sulfate permease family protein	267	51
1793	gi2635979	Bacillus subtilis	similar to transporter	249	46
1793	gi14024597	Mesorhizobium loti	sulfate transporter family protein	231	45
1794	gi1799719	Escherichia coli	similar to	922	100
1794	gi15156677	Agrobacterium tumefaciens	AGR_C_2926p	452	50
1794	gi15074970	Sinorhizobium meliloti	HYPOTHETICAL TRANSMEMBRANE PROTEIN	414	47
1795	gi13325242	Homo sapiens	clone MGC:4033 IMAGE:2820092, mRNA, complete cds.	228	81
1795	gi1644366	Rattus norvegicus	ninjurin1	228	81
1795	gi3077901	Mus musculus	ninjurin	228	81
1796	gi15987491	Homo sapiens	tumor endothelial marker 5 precursor (TEM5) mRNA, complete cds.	5742	89
1796	AAB71863	Homo sapiens	Human h15571 GPCR.	5742	89
1796	gi15987499	Mus musculus	tumor endothelial marker 5 precursor	5030	79
1797	gi13938575	Homo sapiens	Similar to RIKEN cDNA 2610511E22 gene, clone MGC:4251 IMAGE:3028940, mRNA, complete cds.	1331	100
1797	AAY91598	Homo sapiens	Human secreted protein sequence encoded by gene 8 SEQ ID NO:271.	1322	100
1797	gi15029776	Mus musculus	RIKEN cDNA 2610511E22 gene	1317	98
1798	gi42989	Escherichia coli	SecY (PrfA) polypeptide (aa 1-443)	953	100
1798	gi606234	Escherichia coli	secY	953	100
1798	gi15978329	Yersinia pestis	preprotein translocase SecY subunit	941	97
1800	gi10435708	Homo sapiens	cDNA FLJ13639 fis, clone PLACE1011219, weakly similar to PROBABLE OXIDOREDUCTASE (EC 1.-.-.-).	289	84
1800	AAB94698	Homo sapiens	Human protein sequence SEQ ID NO:15680.	289	84

Table 2A  
226

SEQ ID	Hit ID	Species	Description	S score	% Identity
1800	gi15292559	Drosophila melanogaster	SD07613p	101	67
1801	gi7415511	Homo sapiens	mRNA for peptide transporter 3, complete cds.	1281	89
1801	gi13810437	Rattus norvegicus	peptide histidine transporter 1 homolog rPHT2	1066	83
1801	AAY86224	Homo sapiens	Human secreted protein HDPWU34, SEQ ID NO:139.	1280	100
1802	gi387012	Homo sapiens	Human pepsinogen gene, exon 9.	615	97
1802	gi387013	Homo sapiens	Human pepsinogen A (15.0) gene, exon 9, clone cgHGP2.	615	97
1802	AAB66589	Homo sapiens	Human pepsin.	615	97
1803	gi15680159	Homo sapiens	Similar to claudin 2, clone MGC:20191 IMAGE:4645075, mRNA, complete cds.	1113	99
1803	gi10503980	Homo sapiens	clone SP82 claudin 2 mRNA, complete cds.	1113	99
1803	gi9755009	Homo sapiens	claudin-2 mRNA, complete cds.	1113	99
1804	gi476222	Homo sapiens	Human anion exchanger 3 brain isoform (bAE3) mRNA, complete cds.	317	74
1804	gi10953762	Mus musculus	anion exchanger 3 cardiac isoform	317	74
1804	gi309095	Mus musculus	AE3 protein	317	74
1805	AAB29632	Homo sapiens	Human pollinosis-associated gene 581-encoded protein, SEQ ID NO:12.	2518	55
1805	AAY70023	Homo sapiens	Human Protease and associated protein-17 (PPRG-17).	2068	58
1805	gi13529590	Mus musculus	Similar to ubiquitin specific protease 20	1924	56
1806	AAY99363	Homo sapiens	Human PRO1380 (UNQ717) amino acid sequence SEQ ID NO:79.	1773	98
1806	gi12656639	Homo sapiens	equilibrative nucleoside transporter 3 (ENT3) mRNA, complete cds.	1767	98
1806	AAY82285	Homo sapiens	Human ENT1 receptor SEQ ID NO:1.	1767	98
1807	AAG81410	Homo sapiens	Human AFP protein sequence SEQ ID NO:338.	155	27
1807	gi9623190	Mus musculus	sphingosine-1-phosphate phosphohydrolase	159	30
1807	gi15778670	Mus musculus	sphingosine-1-phosphate phosphatase	159	30
1808	gi15082281	Homo sapiens	Similar to steroid dehydrogenase homolog, clone MGC:10252 IMAGE:3622879, mRNA, complete cds.	231	93
1808	gi15214803	Homo sapiens	Similar to steroid dehydrogenase homolog, clone MGC:13329 IMAGE:4281565, mRNA, complete cds.	231	93
1808	gi5531815	Homo sapiens	steroid dehydrogenase homolog	231	93
1810	gi14270513	Homo sapiens	partial SLC22A3 gene for organic cation transporter 3, exon 1 and joined CDS.	775	99
1810	gi3581982	Homo sapiens	mRNA for extraneuronal	775	99

Table 2A  
227

SEQ ID	Hit ID	Species	Description	S score	% Identity
			monoamine transporter.		
1810	gi4454795	Mus musculus	organic cation transporter 3	720	89
1811	gi1212965	Homo sapiens	H.sapiens mRNA for transmembrane protein rnp24.	266	91
1811	AAB74750	Homo sapiens	Human secreted protein sequence encoded by gene 18 SEQ ID NO:59.	266	91
1811	AAB74759	Homo sapiens	Human secreted protein sequence encoded by gene 18 SEQ ID NO:68.	266	91
1812	gi15082375	Homo sapiens	Similar to transmembrane 7 superfamily member 1 (upregulated in kidney), clone MGC:20076 IMAGE:4643216, mRNA, complete cds.	711	98
1812	gi13096836	Mus musculus	Similar to transmembrane 7 superfamily member 1 (upregulated in kidney)	690	95
1812	AAV69992	Homo sapiens	Human receptor-associated protein from Incyte clone 786873.	412	72
1813	gi12653811	Homo sapiens	zinc finger protein 219, clone MGC:1124 IMAGE:3347777, mRNA, complete cds.	500	38
1813	gi6899807	Homo sapiens	mRNA for zinc finger protein 219, complete cds.	496	38
1813	gi14549186	Mus musculus	zinc finger protein 219	490	36
1814	gi4324468	Homo sapiens	LAG1 protein (LAG1) gene, exon 7 and complete cds.	1385	93
1814	gi183051	Homo sapiens	Human growth/differentiation factor 1 (GDF-1) mRNA, complete cds.	1385	93
1814	AAR20230	Homo sapiens	hUOG-1.	1385	93
1815	gi6009515	Xenopus laevis	epidermis specific serine protease	338	41
1815	gi9757698	Xenopus laevis	embryonic serine protease-1	314	46
1815	gi13277969	Mus musculus	Similar to protease, serine, 8 (prostasin)	310	38
1816	gi41875	Escherichia coli	KefC potassium efflux system	706	100
1816	gi216472	Escherichia coli	kefc potassium efflux system	706	100
1816	gi1786232	Escherichia coli K12	K+ efflux antiporter, glutathione-regulated	706	100
1817	gi1054578	Escherichia coli	dioxygenase	1510	100
1817	gi1786565	Escherichia coli K12	taurine dioxygenase, 2-oxoglutarate-dependent	1510	100
1817	gi13359879	Escherichia coli O157:H7	taurine dioxygenase	1506	99
1818	gi1279401	Escherichia coli	SapB protein	928	98
1818	gi1742115	Escherichia coli	Peptide transport system permease protein SapB.	928	98
1818	gi13361335	Escherichia coli O157:H7	homolog of Salmonella peptide transport permease protein	928	98
1819	gi9652147	Homo sapiens	transmembrane-type protein tyrosine phosphatase H (PTPRH) gene, exon 20 and complete cds.	4639	100
1819	gi475004	Homo sapiens	mRNA for protein tyrosine phosphatase precursor, complete cds.	4308	93

Table 2A  
228

SEQ ID	Hit ID	Species	Description	S score	% Identity
1819	gi1321659	Rattus norvegicus	brain-enriched membrane-associated protein tyrosine phosphatase (BEM)-2	920	77
1820	gi7141125	Homo sapiens	tissue-type heart Ellis-van Creveld syndrome protein (EVC) mRNA, complete cds.	1277	100
1820	gi7141127	Homo sapiens	tissue-type brain Ellis-van Creveld syndrome protein (EVC) mRNA, complete cds.	1277	100
1820	gi7271903	Homo sapiens	DWF-1 mRNA, complete cds.	1271	99
1821	AAG71453	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1134.	568	98
1821	AAG72370	Homo sapiens	Human OR-like polypeptide query sequence, SEQ ID NO: 2051.	568	98
1821	AAE04556	Homo sapiens	Human G-protein coupled receptor-12 (GCREC-12) protein.	558	100
1822	gi10440040	Homo sapiens	cDNA: FLJ23375 fis, clone HEP16206.	757	98
1822	AAB95094	Homo sapiens	Human protein sequence SEQ ID NO:17042.	98	43
1822	gi6002197	Homo sapiens	H.sapiens NDUFV3 gene, exon 1.	78	34

Table 2B  
229

SEQ ID	Hit ID	Species	Description	S score	% Identity
912	gi601918	Homo sapiens	glutathione S-transferase theta 2	1263	100
912	gi769703	Rattus norvegicus	glutathione S-transferase subunit Yrs	995	78
912	gi220757	Rattus norvegicus	glutathione S-transferase Yrs-Yrs	995	78
913	gi13872813	Homo sapiens	fibulin-6	4538	93
913	gi14575679	Homo sapiens	hemicentin	4533	93
913	ABB10492	Homo sapiens	Human cDNA SEQ ID NO: 800.	3343	90
914	AAU80378	Homo sapiens	Human DOPD-like protein NOV1.	465	98
914	AAR83048	Homo sapiens	Human macrophage migration inhibitory factor-3 (MIF-3).	450	94
914	gi2104581	Homo sapiens	phenylpyruvate tautomerase II	450	94
915	AAR83048	Homo sapiens	Human macrophage migration inhibitory factor-3 (MIF-3).	546	86
915	gi2104581	Homo sapiens	phenylpyruvate tautomerase II	546	86
915	gi3047378	Homo sapiens	D-dopachrome tautomerase	546	86
916	AAR89197	Homo sapiens	Human hepatocellular growth factor single chain precursor protein.	3466	91
916	gi219681	Homo sapiens	HGF activator precursor	3466	91
916	gi4190954	Homo sapiens	hepatocyte growth factor activator	3466	91
917	gi5441937	Homo sapiens	laminin beta precursor; similar to AAB92586 (PID:g2708707)	8964	100
917	AAV15457	Homo sapiens	Human laminin beta 4 protein.	8838	93
917	AAV15459	Homo sapiens	SEQ ID 5 of WO9919347.	6042	97
918	AAG65887	Homo sapiens	Amino acid sequence of GSK gene Id 14936.	5050	100
918	AAG65888	Homo sapiens	Amino acid sequence of GSK gene Id 14936.	4582	84
918	AAG68261	Homo sapiens	Human POLY8 protein sequence SEQ ID NO:16.	4543	83
919	ABB74955	Homo sapiens	Human lung tumour L524S variant protein sequence SEQ ID NO:166.	905	100
919	ABB74954	Homo sapiens	Human lung tumour L524S variant protein sequence SEQ ID NO:165.	905	100
919	AAE23751	Homo sapiens	Human parathyroid related peptide, PTHrP (1-173).	905	100
920	gi20269129	Homo sapiens	MEGF6	2584	95
920	AAE17919	Homo sapiens	Human gene 1 encoded serine protease, HMGBM65.	2468	99
920	AAV72091	Homo sapiens	Human serine protease #2 encoded by clone HMGBM65.	2468	99
921	ABB90774	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 281.	346	100
921	ABB90773	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 279.	346	100
921	ABB90772	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 277.	346	100
922	AAU97039	Homo sapiens	Human LP protein LP190.	2054	89
922	ABG30500	Homo sapiens	Human Carboxypeptidase A.	2054	89
922	AAB47565	Homo sapiens	Protease PR7S-7.	2054	89
923	AAM52240	Homo sapiens	Human MFAP4 SEQ ID NO 3.	1198	99
923	gi790817	Homo sapiens	microfibril-associated glycoprotein 4	1198	99
923	AAM52239	Homo sapiens	Human MAG4V SEQ ID NO 1.	1197	100
924	AAU81960	Homo sapiens	Human PRO536.	1648	100
924	AAB65173	Homo sapiens	Human PRO536 (UNQ337) protein	1648	100

Table 2B  
230

SEQ ID	Hit ID	Species	Description	S score	% Identity
			sequence SEQ ID NO:97.		
924	AAB94830	Homo sapiens	Human protein sequence SEQ ID NO:15991.	1648	100
925	AAU10497	Homo sapiens	Human Apolipoprotein A-II, APOA2.	438	90
925	gi13528981	Homo sapiens	apolipoprotein A-II	438	90
925	gi21429231	Homo sapiens	apolipoprotein AII	438	90
926	AAAY76156	Homo sapiens	Human secreted protein encoded by gene 33.	419	90
927	gi13097252	Homo sapiens	Similar to FK506 binding protein 2 (13 kDa)	648	91
927	gi337370	Homo sapiens	rapamycin- and FK506-binding protein	648	91
927	AAR93551	Homo sapiens	Human FKBP-13 immunophilin protein.	628	90
928	AAG73789	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:4553.	230	95
928	AAB53360	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:900.	230	95
930	gi17982522	Brucella melitensis	(S)-2-hydroxy-acid oxidase chain D	1151	49
930	gi17740510	Agrobacterium tumefaciens str. C58 (U. Washington)	FAD dependent oxidoreductase	1149	49
930	gi15157181	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_C_3718p	1149	49
931	ABB89770	Homo sapiens	Human polypeptide SEQ ID NO 2146.	366	70
931	AAB75367	Homo sapiens	Human secreted protein #26.	366	70
931	AAU04353	Homo sapiens	Mammalian toxicological response marker protein #5.	366	70
932	ABB55723	Homo sapiens	Human polypeptide SEQ ID NO 52.	1107	99
932	AAU39014	Homo sapiens	Human secreted protein yc2 1.	1107	99
932	AAM94621	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 3279.	203	67
933	gi16648246	Drosophila melanogaster	GH27263p	228	33
933	gi1711197	Xenopus laevis	Xfringe2	184	25
933	gi1679784	Xenopus laevis	radical fringe	184	26
934	AAAY73383	Homo sapiens	HTRM clone 2280456 protein sequence.	1571	100
934	AAG75282	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6046.	619	99
934	AAM41045	Homo sapiens	Human polypeptide SEQ ID NO 5976.	89	29
935	AAG65916	Homo sapiens	Amino acid sequence of GSK gene Id 239881.	3041	99
935	gi15705411	Homo sapiens	peptidoglycan recognition protein L precursor	3041	99
935	AAG65915	Homo sapiens	Amino acid sequence of GSK gene Id 239881.	2892	99
936	AAU97218	Homo sapiens	Human G protein-coupled receptor, TGR25.	850	99

Table 2B  
231

SEQ ID	Hit ID	Species	Description	S score	% Identity
936	AAE23415	Homo sapiens	Human G-protein coupled receptor-7 (GCREC-7).	850	99
936	ABB12463	Homo sapiens	Human bone marrow expressed protein SEQ ID NO: 302.	769	100
937	gi1711232	Homo sapiens	14-3-3 protein eta chain	1094	94
937	gi460779	Homo sapiens	14-3-3 eta subtype	1094	94
937	gi1477931	Homo sapiens	14.3.3 eta chain	1094	94
938	gi16118441	Oryctolagus cuniculus	S-100 calcium-binding protein beta subunit	479	100
938	gi12804681	Homo sapiens	S100 calcium-binding protein, beta (neural)	479	100
938	gi337730	Homo sapiens	S100 protein beta subunit	479	100
939	AAW75082	Homo sapiens	Human secreted protein encoded by gene 26 clone HTLEV12.	392	92
939	AAM95398	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 4056.	71	50
939	AAU29184	Homo sapiens	Human PRO polypeptide sequence #161.	71	50
940	AAU29315	Homo sapiens	Human PRO polypeptide sequence #292.	3743	99
940	ABB75753	Homo sapiens	Human pancreas GP354.	3104	100
940	ABB75751	Homo sapiens	Human immunoglobulin superfamily member GP354.	3073	99
941	AAO21477	Homo sapiens	Human Ngr2 protein sequence.	2133	100
941	AAO21482	Homo sapiens	Mature human Ngr2 protein sequence.	2082	100
941	gi20987877	Mus musculus	similar to Nogo receptor	885	58
942	AAB25674	Homo sapiens	Human secreted protein sequence encoded by gene 10 SEQ ID NO:63.	768	98
942	AAB36613	Homo sapiens	Human FLEXHT-35 protein sequence SEQ ID NO:35.	238	90
942	gi14603247	Homo sapiens	Similar to RIKEN cDNA 5730409G15 gene	238	90
943	AAE22093	Homo sapiens	Human kidney specific renal cell carcinoma (KSRCC) protein.	851	98
943	AAY85678	Homo sapiens	Human kidney disease associated protein SEQ ID 10.	751	98
943	gi3127193	Rattus norvegicus	kidney-specific protein	686	75
944	ABP41513	Homo sapiens	Human ovarian antigen HCOOX52, SEQ ID NO:2645.	2558	100
944	gi1197499	Homo sapiens	C1 inhibitor	2527	100
944	AAW18207	Homo sapiens	Wild-type C1 inhibitor.	2524	99
945	gi17982052	Brucella melitensis	RIBOSOMAL PROTEIN L11 METHYLTRANSFERASE	241	36
945	gi15157854	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_C_4799p	207	36
945	gi18145768	Clostridium perfringens str. 13	probable ribosomal protein L11 methyltransferase	118	37
946	AAB03948	Homo sapiens	Human mesenchymal stem cell polypeptide.	462	98
946	AAB64909	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:87.	165	63

Table 2B

232

SEQ ID	Hit ID	Species	Description	S score	% Identity
946	AAU01649	Homo sapiens	Human secreted protein immunogenic epitope encoded by gene #12.	165	63
947	AAE13809	Homo sapiens	Human lung tumour-specific protein SAL-82.	2906	99
947	AAB44467	Homo sapiens	Human lung tumour-specific antigen encoded by cDNA #103.	2906	99
947	gi13958036	Homo sapiens	RUFY1	2906	99
949	AAB88326	Homo sapiens	Human membrane or secretory protein clone PSEC0021.	541	74
949	AAB92475	Homo sapiens	Human protein sequence SEQ ID NO:10549.	541	74
949	AAE03226	Homo sapiens	Human gene 7 encoded secreted protein HNTDL21, SEQ ID NO:76.	541	74
950	ABP47877	Homo sapiens	Human polypeptide SEQ ID NO 307.	1303	100
950	ABP47870	Homo sapiens	Human polypeptide SEQ ID NO 300.	959	99
950	ABP48038	Homo sapiens	Human polypeptide SEQ ID NO 468.	835	98
951	AAU12382	Homo sapiens	Human PRO792 polypeptide sequence.	1138	87
951	AAB24416	Homo sapiens	Human PRO792 protein sequence SEQ ID NO:155.	1138	87
951	AAB24055	Homo sapiens	Human PRO792 protein sequence SEQ ID NO:31.	1138	87
952	gi7670746	Homo sapiens	UDP-glucose:glycoprotein glucosyltransferase 1 precursor	7974	98
952	gi13275621	synthetic construct	Rat RUGT	7317	91
952	gi7677176	Rattus norvegicus	UDP-glucose glycoprotein:glucosyltransferase precursor	7317	91
953	gi20810499	Homo sapiens	Similar to RIKEN cDNA 2900074C18 gene	1237	100
953	AAM42005	Homo sapiens	Human polypeptide SEQ ID NO 6936.	404	48
953	AAB64390	Homo sapiens	Amino acid sequence of human intracellular signalling molecule INTRA22.	402	50
954	gi18676660	Homo sapiens	FLJ00229 protein	2292	98
954	AAM40342	Homo sapiens	Human polypeptide SEQ ID NO 3487.	1454	60
954	AAM42128	Homo sapiens	Human polypeptide SEQ ID NO 7059.	765	60
955	gi21707216	Homo sapiens	LOC146556	2100	99
955	AAG81399	Homo sapiens	Human AFP protein sequence SEQ ID NO:316.	1910	100
955	AAB61421	Homo sapiens	Human TANGO 300 protein.	1904	99
956	gi21667020	Homo sapiens	mutant I beta-1,6-N-acetylglucosaminyltransferase C form	1637	99
956	gi21667011	Homo sapiens	I beta-1,6-N-acetylglucosaminyltransferase C form	1637	99
956	gi21667015	Homo sapiens	mutant I beta-1,6-N-acetylglucosaminyltransferase C form	1633	99



Table 2B

233

SEQ ID	Hit ID	Species	Description	S score	% Identity
957	gi2065165	Homo sapiens	collagen type XIV	773	93
957	AAG00322	Homo sapiens	Human secreted protein, SEQ ID NO: 4403.	482	98
957	gi288875	Gallus gallus	collagen XIV	461	63
958	AAG75509	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6273.	2482	94
958	AAB84606	Homo sapiens	Amino acid sequence of matrix metalloproteinase collagenase 1.	2482	94
958	AAE10415	Homo sapiens	Human matrix metalloproteinase-1 (MMP-1) protein.	2482	94
959	AAAY25868	Homo sapiens	Human secreted protein fragment encoded from gene 57.	993	99
959	gi15145797	Sus scrofa	basic proline-rich protein	88	42
959	gi3413810	Mus musculus	Bassoon	84	31
960	AAG77807	Homo sapiens	Human HSIFL59 serine/threonine phosphatase protein sequence.	473	98
960	AAB58288	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 626.	473	98
960	AAAY59744	Homo sapiens	Human normal ovarian tissue derived protein 21.	473	98
961	gi4019247	Ateline herpesvirus 3	thymidine kinase	71	46
962	gi9956936	Mus musculus	Su(var)3-9 homolog Suv39h2	1310	86
962	AAM78701	Homo sapiens	Human protein SEQ ID NO 1363.	1236	100
962	AAW05261	Homo sapiens	Chromatin regulator protein SUV39H.	814	55
963	AAG03840	Homo sapiens	Human secreted protein, SEQ ID NO: 7921.	439	80
963	gi1698653	Homo sapiens	NADPH:ubiquinone oxidoreductase subunit B13	439	80
963	gi12654023	Homo sapiens	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5 (13kD, B13)	439	80
964	gi37347	Homo sapiens	TRG gamma chain (AA 1-340)	1657	94
964	gi339407	Homo sapiens	Ti antigen CD3-associated protein precursor	1627	93
964	gi37018	Homo sapiens	pre-gamma-chain (AA -14 to 309)	1436	87
965	AAE06606	Homo sapiens	Human protein having hydrophobic domain, HP10794.	566	96
965	AAM94615	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 3273.	566	96
965	gi18490535	Mus musculus	RIKEN cDNA 2610528J11 gene	388	78
966	gi13654639	Bos taurus	D-glucuronyl C5 epimerase	3159	97
966	gi13442978	Mus musculus	D-glucuronyl C5-epimerase	3139	95
966	gi11935177	Mus musculus	heparin/heparan sulfate:glucuronic acid C5 epimerase	3134	95
967	AAG72204	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1885.	1167	77
967	AAG71875	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1556.	1138	74
967	AAG71816	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1497.	1102	72
968	AAB73679	Homo sapiens	Human oxidoreductase protein ORP-12.	1918	100
968	gi12655133	Homo sapiens	Similar to CGI-63 protein	1918	100
968	ABB89788	Homo sapiens	Human polypeptide SEQ ID NO	1907	99

Table 2B  
234

SEQ ID	Hit ID	Species	Description	S score	% Identity
			2164.		
969	AAY28995	Homo sapiens	Tumour suppressor Del-27 protein sequence.	2309	59
969	AAY15344	Homo sapiens	Tumour suppressor protein del-27.	2309	59
969	gi6062874	Homo sapiens	candidate tumor suppressor protein DICE1	2309	59
970	gi387011	Homo sapiens	pyruvate dehydrogenase E1-alpha precursor	2187	99
970	gi35381	Homo sapiens	pyruvate dehydrogenase E1' subunit (AA 1 - 390)	2049	100
970	gi219982	Homo sapiens	alpha subunit	2049	100
971	gi2275569	Homo sapiens	TCRBV23S1	660	99
971	gi2104755	Homo sapiens	T cell receptor V-beta 23	660	99
971	gi495404	Pan troglodytes	T cell receptor beta chain	657	88
972	AAM95505	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 4163.	1469	99
972	ABB96188	Homo sapiens	Human testicular antigen SEQ ID NO: 1572.	1469	99
972	AAB75360	Homo sapiens	Human secreted protein #19.	1272	100
973	AAE02937	Homo sapiens	Human TFRP protein.	3040	94
973	gi4996563	Homo sapiens	protein inhibitor of activated STAT3	3040	94
973	gi12654633	Homo sapiens	protein inhibitor of activated STAT3	3040	94
974	gi15099957	Homo sapiens	diacylglycerol acyltransferase 2-like protein	855	93
974	gi15099955	Mus musculus	diacylglycerol acyltransferase 2-like protein	675	59
974	AAY94889	Homo sapiens	Human protein clone HP02485.	576	56
975	AAB92669	Homo sapiens	Human protein sequence SEQ ID NO:11033.	1432	100
975	ABG40833	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 30498.	399	98
975	AAM31319	Homo sapiens	Peptide #5356 encoded by probe for measuring placental gene expression.	399	98
976	gi4210474	Navel orange infectious mottling virus	polyprotein	71	21
977	gi1209685	Homo sapiens	salivary peroxidase	3754	96
977	gi163307	Bos taurus	lactoperoxidase	3166	80
977	gi11990122	Camelus dromedarius	peroxidase	3153	81
978	AAY01604	Homo sapiens	Amino acid sequence of the human defensin (Def-X) protein.	501	97
978	gi29735	Homo sapiens	corticostatin/defensin HP-4 precursor	214	46
978	AAR15222	Homo sapiens	Chronic myelogenous leukaemia-derived myeloid-related protein.	211	47
979	AAG81415	Homo sapiens	Human AFP protein sequence SEQ ID NO:348.	848	100
980	gi6633820	Arabidopsis thaliana	FIN19.20	79	26
980	gi6180001	Capra hircus	pregnancy-associated glycoprotein-8	73	30
980	gi20071290	Mus musculus	Similar to solute carrier family 26, member 7	71	25
981	gi4309953	Homo sapiens	T cell receptor gamma chain; similar to PID:g339160	951	95

Table 2B

235

SEQ ID	Hit ID	Species	Description	S score	% Identity
981	gi4309950	Homo sapiens	T cell receptor gamma chain; match to S08328 (PID:g106470)	917	88
981	ABG39493	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 29158.	679	77
982	gi1223888	synthetic construct	T cell receptor alpha chain	1066	79
982	gi338766	Homo sapiens	T-cell receptor precursor	987	73
982	gi3089419	Homo sapiens	T cell receptor alpha chain	974	71
983	ABB07518	Homo sapiens	Human drug metabolizing enzyme (DME) (ID: 7474438CD1).	418	29
983	AAB73512	Homo sapiens	Human transferase HTFS-19, SEQ ID NO:19.	409	30
983	gi14249942	Homo sapiens	Similar to RIKEN cDNA 0610008P16 gene	409	30
984	AAG71286	Homo sapiens	Human gene 9-encoded secreted protein HMSDL37, SEQ ID NO:135.	311	90
984	AAG71251	Homo sapiens	Human gene 9-encoded secreted protein HMSDL37, SEQ ID NO:99.	311	90
984	gi13096922	Mus musculus	Similar to nadrin	74	44
985	gi4519541	Mus musculus	thrombospondin type 1 domain	1296	88
985	ABP61846	Homo sapiens	Human polypeptide SEQ ID NO 200.	644	46
985	AAB99220	Homo sapiens	Human thrombospondin-30.	644	46
986	AAG81417	Homo sapiens	Human AFP protein sequence SEQ ID NO:352.	287	100
987	ABB53264	Homo sapiens	Human polypeptide #4.	1904	100
987	gi18565270	Homo sapiens	Lib	1904	100
987	gi18565266	Rattus norvegicus	Lib	1621	83
988	AAB59022	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 730.	1385	100
988	gi14124958	Homo sapiens	nucleotide binding protein 2 (E.coli MinD like)	1385	100
988	gi13559170	Homo sapiens	C447E6.1 (nucleotide binding protein 1 (E.coli MinD like) )	1380	100
989	AAU99292	Homo sapiens	Human chordin-like associated protein #1.	3916	100
989	AAU99293	Homo sapiens	Human chordin-like associated protein #2.	3073	81
989	AAE07119	Homo sapiens	Human gene 12 encoded secreted protein fragment, SEQ ID NO:136.	2411	98
990	AAB26105	Homo sapiens	Human DAN/Cerberus-related protein 6 (hDCR6) #1.	1693	87
990	AAE17089	Homo sapiens	Human osteolevin protein.	439	66
990	ABB07209	Homo sapiens	Human cloaked-2 polypeptide sequence.	439	66
991	gi9964099	Chlamydia trachomatis	inclusion membrane protein	74	35
992	gi21619848	Homo sapiens	Similar to immunoglobulin lambda joining 3	1002	89
992	gi4490538	Homo sapiens	lambda-immunoglobulin light chain	927	81
992	gi33746	Homo sapiens	immunoglobulin lambda light chain	914	80
993	AAG03466	Homo sapiens	Human secreted protein, SEQ ID NO: 7547.	97	100

Table 2B

236

SEQ ID	Hit ID	Species	Description	S score	% Identity
993	gi4063766	Aspergillus nidulans	chitinase	91	30
993	gi3617766	Anopheles gambiae	ICHIT protein	88	27
994	gi203246	Rattus norvegicus	cell adhesion-like molecule	1739	97
994	gi514374	Homo sapiens	opioid-binding cell adhesion molecule	1665	97
994	gi586	Bos taurus	put. pre-OPCAM (AA 1 - 345)	1640	93
995	AAB53088	Homo sapiens	Human angiogenesis-associated protein PRO328, SEQ ID NO:132.	2391	91
995	AAB80260	Homo sapiens	Human PRO328 protein.	2391	91
995	AAU12351	Homo sapiens	Human PRO328 polypeptide sequence.	2391	91
996	AAB85144	Homo sapiens	Human NKCR polypeptide (clone ID HMSOM53).	1205	83
996	AAAY96226	Homo sapiens	Human high affinity Fc receptor, FcgammaRI.	354	41
996	gi31332	Homo sapiens	FcRI (AA 1-374)	354	41
997	ABB50835	Homo sapiens	Human secreted protein encoded by gene 80 SEQ ID NO:788.	74	36
998	ABP61434	Homo sapiens	Human NF-kB activating protein SEQ ID NO 21.	417	100
998	AAU07442	Homo sapiens	Human Wnt1 Upregulated protein 2 (WUP2).	417	100
998	AAU07441	Homo sapiens	Human Wnt1 Upregulated protein 1 (WUP1).	417	100
999	AAB08732	Homo sapiens	Amino acid sequence of a human OLD-35 polypeptide.	1671	85
999	gi20372922	Homo sapiens	polynucleotide phosphorylase-like protein	1671	85
999	AAB92684	Homo sapiens	Human protein sequence SEQ ID NO:11065.	1265	88
1000	gi199582	Mus musculus	B(2)-microglobulin	616	95
1000	gi50105	Mus musculus	beta2-microglobulin precursor (aa - 20 to 99)	614	95
1000	gi199576	Mus musculus	B(2)-microglobulin	609	94
1001	AAB93335	Homo sapiens	Human protein sequence SEQ ID NO:12441.	672	99
1001	AAE21620	Homo sapiens	Human gene 7 encoded secreted protein, SEQ ID NO:92.	427	91
1001	AAE21604	Homo sapiens	Human gene 7 encoded secreted protein HTFOE85, SEQ ID NO:76.	402	90
1002	gi15559608	Homo sapiens	Similar to zinc finger protein 16 (KOX 9)	2079	100
1002	gi488555	Homo sapiens	zinc finger protein ZNF135	749	65
1002	AAB21006	Homo sapiens	Human nucleic acid-binding protein, NuABP-10.	747	54
1003	ABB79480	Homo sapiens	Human zinc finger protein 75.68.	1689	56
1003	gi1769491	Homo sapiens	kruppel-related zinc finger protein	1672	53
1003	AAM39130	Homo sapiens	Human polypeptide SEQ ID NO 2275.	1669	52
1005	AAB23641	Homo sapiens	Human secreted protein SEQ ID NO: 97.	609	100
1005	gi13129458	Oryza sativa	polyprotein	89	31

Table 2B  
237

SEQ ID	Hit ID	Species	Description	S score	% Identity
		[Oryza sativa (japonica cultivar-group)]			
1005	gi7228457	Oryza sativa (japonica cultivar-group)	Similar to Sorghum bicolor Gypsy-Ty3 type retrotransposon RetroSor1, polyprotein. (AF098806)	87	29
1006	gi11493473	Homo sapiens	PRO2225	163	67
1007	AAM52305	Homo sapiens	Human zyxine.	2815	91
1007	AAG68191	Homo sapiens	Zyxin protein SEQ ID NO:107.	2815	91
1007	gi1545954	Homo sapiens	zyxin	2815	91
1008	AAM00955	Homo sapiens	Human bone marrow protein, SEQ ID NO: 431.	704	100
1008	gi18028488	Homo sapiens	cytosolic leucine-rich protein	631	99
1008	gi21666364	Bos taurus	leucine-rich protein	588	90
1009	AAE01420	Homo sapiens	Human secreted protein fragment, SEQ ID NO:144.	612	100
1009	gi17225457	Homo sapiens	autism-related protein 1	76	38
1009	AAM79126	Homo sapiens	Human protein SEQ ID NO 1788.	72	29
1010	gi16877231	Homo sapiens	Similar to RIKEN cDNA 2700019D07 gene	997	100
1010	AAM39593	Homo sapiens	Human polypeptide SEQ ID NO 2738.	94	100
1010	gi7302568	Drosophila melanogaster	CG15073-PA	91	19
1011	AAB43434	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:879.	413	78
1011	AAAY07039	Homo sapiens	Breast cancer associated antigen precursor sequence.	413	78
1011	gi17932966	Homo sapiens	ribosomal protein P1	413	78
1012	AAW88457	Homo sapiens	Human lysophospholipase IHLP.	1125	93
1012	ABP51416	Homo sapiens	Human MDDT SEQ ID NO 438.	843	100
1012	AAB75386	Homo sapiens	Human secreted protein #45.	825	100
1013	AAG81374	Homo sapiens	Human AFP protein sequence SEQ ID NO:266.	919	98
1013	gi20977549	Danio rerio	DT1P1A10-like protein	354	41
1013	gi665970	Saccharomyces cerevisiae	Ylr435wp	169	26
1014	gi20988991	Mus musculus	RIKEN cDNA 2810405K02 gene	920	88
1014	AAM93895	Homo sapiens	Human polypeptide, SEQ ID NO: 4031.	777	100
1014	ABB84903	Homo sapiens	Human PRO1198 protein sequence SEQ ID NO:174.	187	36
1015	AAB56791	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1369.	496	98
1015	gi19401678	Giardia intestinalis	endosomal AAA ATPase-like protein	76	38
1016	ABB44579	Homo sapiens	Human wound healing related polypeptide SEQ ID NO 36.	655	100
1016	AAB53427	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:967.	655	100
1016	gi2286227	Bos taurus	myocardial vascular inhibition factor	655	100
1017	AAB81188	Homo sapiens	Human zinc finger protein 52 (ZFP-52).	2389	98
1017	AAB95368	Homo sapiens	Human protein sequence SEQ ID NO:17684.	1712	100

Table 2B  
238

SEQ ID	Hit ID	Species	Description	S score	% Identity
1017	ABP51456	Homo sapiens	Human MDDT SEQ ID NO 478.	1561	99
1018	gi6434346	Caenorhabditis elegans	Y105E8B.4	77	33
1019	AAG66831	Homo sapiens	Human DNA-dependent protein kinase 9.	443	98
1019	AAB90816	Homo sapiens	Human shear stress-response protein SEQ ID NO: 140.	365	100
1019	AAG66832	Homo sapiens	Human DNA-dependent protein kinase 9 N-terminal peptide.	79	100
1020	gi1377897	Homo sapiens	heart protein	1631	100
1020	gi7209525	Homo sapiens	DRAL/Slim3/FHL2	1625	99
1020	gi5825391	Mus musculus	four and half LIM domain protein 2	1524	91
1021	gi18139947	Homo sapiens	HLCDGP1	893	99
1021	gi7770259	Homo sapiens	PRO2975	794	100
1021	gi1237130	Escherichia coli	O antigen polymerase	85	26
1022	gi19263712	Homo sapiens	Similar to LOC146557	972	100
1022	AAG81348	Homo sapiens	Human AFP protein sequence SEQ ID NO:214.	528	62
1022	AAM88837	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:16430.	179	80
1023	AAG01390	Homo sapiens	Human secreted protein, SEQ ID NO: 5471.	297	100
1023	gi456681	Pseudorabies virus	helicase	80	25
1023	AAM85692	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:13285.	78	63
1024	gi4235144	Homo sapiens	BC39498_1	1423	62
1024	gi21265141	Homo sapiens	Similar to zinc finger protein 91 (HPF7, HTF10)	1404	60
1024	gi14348591	Homo sapiens	KRAB zinc finger protein	1403	58
1025	gi18490643	Homo sapiens	Similar to recombination activating gene 2	2849	99
1025	gi165680	Oryctolagus cuniculus	recombination activating protein	2666	91
1025	gi2576246	Mus musculus	RAG-2 protein	2594	88
1026	AAG75278	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6042.	1058	95
1026	gi600255	Gallus gallus	caldesmon	147	28
1026	gi211896	Gallus gallus	h-caldesmon	147	28
1027	AAV87341	Homo sapiens	Human signal peptide containing protein HSPP-118 SEQ ID NO:118.	699	99
1027	gi12311853	Leishmania major	possible surface antigen	77	31
1028	AAM92844	Homo sapiens	Human digestive system antigen SEQ ID NO: 2193.	86	38
1028	ABB11195	Homo sapiens	Human transmembrane protein homologue, SEQ ID NO:1565.	82	25
1028	ABG66815	Homo sapiens	Human prostate specific protein DEX0283_123.	75	60
1029	AAE02058	Homo sapiens	Human four disulfide core domain (FDCD)-containing protein.	587	41
1029	AAM79986	Homo sapiens	Human protein SEQ ID NO 3632.	579	41
1029	AAM79002	Homo sapiens	Human protein SEQ ID NO 1664.	579	41
1031	ABP42929	Homo sapiens	Human ovarian antigen HPDRS87, SEQ ID NO:4061.	820	93

Table 2B  
239

SEQ ID	Hit ID	Species	Description	S score	% Identity
1031	ABP41129	Homo sapiens	Human ovarian antigen HE2RG21, SEQ ID NO:2261.	820	93
1031	AAU01195	Homo sapiens	Human cyclophilin A protein.	820	93
1032	AAE06643	Homo sapiens	Human G-protein coupled receptor (NGPCR) #2.	671	99
1032	AAE06642	Homo sapiens	Human G-protein coupled receptor (NGPCR) #1.	671	99
1032	gi7291589	Drosophila melanogaster	CG18679-PA	214	34
1033	AAG03055	Homo sapiens	Human secreted protein, SEQ ID NO: 7136.	269	100
1034	AAB38043	Homo sapiens	Fragment of human secreted protein encoded by gene 10 clone HWHGP71.	125	36
1034	gi5305335	Mycobacterium tuberculosis	proline-rich mucin homolog	105	33
1034	gi5917666	Zea mays	extensin-like protein	104	37
1035	AAM93942	Homo sapiens	Human polypeptide, SEQ ID NO: 4126.	3226	99
1035	ABB11422	Homo sapiens	Human Zn finger protein homologue, SEQ ID NO:1792.	2760	96
1035	gi6467206	Homo sapiens	gonadotropin inducible transcription repressor-4	2032	57
1036	AAB95007	Homo sapiens	Human protein sequence SEQ ID NO:16685.	518	86
1036	gi21410398	Mus musculus	RIKEN cDNA 2610034E13 gene	81	32
1036	gi45906	Proteus vulgaris	hlyC protein (AA 1-54)	72	45
1037	AAAY27616	Homo sapiens	Human secreted protein encoded by gene No. 50.	562	99
1037	gi17902598	Rice black streaked dwarf virus	P6 protein	71	25
1038	gi16588681	Homo sapiens	anion transporter/exchanger-9	4295	95
1038	AAE21166	Homo sapiens	Human TRICH-10 protein.	3612	91
1038	gi13344999	Homo sapiens	solute carrier family 26 member 6	1298	37
1039	AAU12254	Homo sapiens	Human PRO4343 polypeptide sequence.	780	100
1039	AAM40835	Homo sapiens	Human polypeptide SEQ ID NO 5766.	780	100
1039	AAAY76141	Homo sapiens	Human secreted protein encoded by gene 18.	780	100
1040	ABB89694	Homo sapiens	Human polypeptide SEQ ID NO 2070.	622	77
1040	AAAY59672	Homo sapiens	Secreted protein 108-006-5-0-E6-FL.	622	77
1040	AAB94543	Homo sapiens	Human protein sequence SEQ ID NO:15290.	618	76
1041	AAAY92710	Homo sapiens	Human membrane-associated protein Zsig24.	704	97
1041	ABB89722	Homo sapiens	Human polypeptide SEQ ID NO 2098.	566	99
1041	AAAY87250	Homo sapiens	Human signal peptide containing protein HSPP-27 SEQ ID NO:27.	566	99
1042	AAU29316	Homo sapiens	Human PRO polypeptide sequence #293.	2932	99
1042	ABB05749	Homo sapiens	Human G protein-coupled receptor	1591	44

Table 2B  
240

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NOV1a protein SEQ ID NO:2.		
1042	gi14572521	Homo sapiens	NEPH1	1519	51
1043	AAU00688	Homo sapiens	Human CD59 protein.	710	100
1043	AAW26318	Homo sapiens	Human CD59.	710	100
1043	AAR80240	Homo sapiens	Human membrane attack complex inhibition factor.	710	100
1044	gi17390957	Mus musculus	Similar to RIKEN cDNA 2010001E11 gene	1455	74
1044	gi6841140	Homo sapiens	HSPC100	498	100
1044	gi17985273	Brucella melitensis	GLUCOSE/GALACTOSE TRANSPORTER	124	22
1045	AAB56632	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1210.	3377	99
1045	gi13097708	Homo sapiens	ribophorin II	3152	100
1045	gi4730801	Homo sapiens	dJ343K2.2.1 (ribophorin II (isoform 1))	3152	100
1046	AAB70690	Homo sapiens	Human hDPP protein sequence SEQ ID NO:7.	598	100
1046	AAG89279	Homo sapiens	Human secreted protein, SEQ ID NO: 399.	598	100
1046	gi13182757	Homo sapiens	HTPAP	598	100
1047	gi2276448	Homo sapiens	MHC class I HLA-A	1794	93
1047	gi6815812	Homo sapiens	MHC class I antigen heavy chain	1794	93
1047	gi1245460	Homo sapiens	MHC class I HLA-A	1786	92
1048	ABP41629	Homo sapiens	Human ovarian antigen HOOJQ91, SEQ ID NO:2761.	675	90
1048	AAB95392	Homo sapiens	Human protein sequence SEQ ID NO:17743.	564	78
1048	AAM79768	Homo sapiens	Human protein SEQ ID NO 3414.	564	78
1049	gi14017773	Mus musculus	Cg10671-like	1517	96
1049	gi14017764	Mus musculus	CG10671-like	1517	96
1049	ABB89676	Homo sapiens	Human polypeptide SEQ ID NO 2052.	957	89
1050	AAG81431	Homo sapiens	Human AFP protein sequence SEQ ID NO:380.	503	97
1050	AAE23305	Homo sapiens	Human nectin-4 protein #4.	128	32
1050	gi19353148	Mus musculus	Similar to poliovirus receptor-related 4	127	27
1051	gi20072551	Mus musculus	RIKEN cDNA 4930511J11 gene	420	45
1051	gi17974542	Homo sapiens	voltage-dependent calcium channel gamma-8 subunit	147	25
1051	gi12836893	Gallus gallus	IPR328-like protein	147	29
1052	ABB84978	Homo sapiens	Human PRO4430 protein sequence SEQ ID NO:324.	436	70
1052	ABB95584	Homo sapiens	Human angiogenesis related protein PRO4430 SEQ ID NO: 324.	436	70
1052	AAU29273	Homo sapiens	Human PRO polypeptide sequence #250.	436	70
1053	AAB88325	Homo sapiens	Human membrane or secretory protein clone PSEC0020.	912	99
1053	AAB53257	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:797.	859	99
1053	AAAY87264	Homo sapiens	Human signal peptide containing protein HSPP-41 SEQ ID NO:41.	315	63
1054	AAU12201	Homo sapiens	Human PRO1779 polypeptide	1819	54



Table 2B  
241

SEQ ID	Hit ID	Species	Description	S score	% Identity
			sequence.		
1054	AAB25594	Homo sapiens	Protein encoded by human secreted protein gene #1.	1819	54
1054	gi1234787	Xenopus laevis	up-regulated by thyroid hormone in tadpoles; expressed specifically in the tail and only at metamorphosis; membrane bound or extracellular protein; C-terminal basic region	1799	53
1055	AAW55035	Homo sapiens	HPURR amino acid sequence.	2014	100
1055	AAW47066	Homo sapiens	Human brain P2X-1 receptor polypeptide.	2014	100
1055	gi4099139	Homo sapiens	P2X4 purinoreceptor	2014	100
1056	AAE03560	Homo sapiens	Human differentially expressed kidney cDNA 22360 encoded protein.	1255	85
1056	AAM42468	Homo sapiens	Human kidney related polypeptide SEQ ID NO 337.	145	90
1056	AAM99653	Homo sapiens	Human excretory related polypeptide SEQ ID NO 390.	145	90
1057	ABB53267	Homo sapiens	Human polypeptide #7.	3372	98
1057	AAO14449	Homo sapiens	Protein of human Zona Pellucida 1 (Zp1).	3367	98
1057	gi972946	Mus musculus	ZP1 precursor	2216	67
1058	gi15779156	Homo sapiens	Similar to RIKEN cDNA 1810073N04 gene	1858	100
1058	gi13097045	Mus musculus	Similar to RIKEN cDNA 1810073N04 gene	1719	91
1058	AAM79693	Homo sapiens	Human protein SEQ ID NO 3339.	1138	100
1059	AAM79993	Homo sapiens	Human protein SEQ ID NO 3639.	1736	89
1059	AAM79009	Homo sapiens	Human protein SEQ ID NO 1671.	1736	89
1059	ABB12000	Homo sapiens	Human prostaglandin DP receptor homologue, SEQ ID NO:2370.	1736	89
1060	AAU79946	Homo sapiens	Human transporter protein sequence.	2907	99
1060	AAE21181	Homo sapiens	Human TRICH-25 protein.	2672	91
1060	gi2811122	Xenopus laevis	NaDC-2	1742	54
1061	AAM79483	Homo sapiens	Human protein SEQ ID NO 3129.	1698	88
1061	AAM78499	Homo sapiens	Human protein SEQ ID NO 1161.	1698	88
1061	ABB11938	Homo sapiens	Human cystinosin homologue, SEQ ID NO:2308.	1698	88
1062	gi12656590	Danio rerio	P2x purinoceptor subunit 4	72	40
1063	AAG68335	Homo sapiens	Human CSP2 protein SEQ ID NO:4.	1354	99
1063	gi19525540	Homo sapiens	lymphocyte effector toxicity activation ligand	1330	98
1063	AAAY36071	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 456.	1252	92
1064	gi15277509	Homo sapiens	Similar to transmembrane 7 superfamily member 2	1548	97
1064	ABP41089	Homo sapiens	Human ovarian antigen HSLGG58, SEQ ID NO:2221.	1501	100
1064	gi18138238	Bos taurus	C-14 sterol reductase	1376	87
1065	AAM93346	Homo sapiens	Human polypeptide, SEQ ID NO: 2891.	5017	98
1065	AAM93761	Homo sapiens	Human polypeptide, SEQ ID NO: 3754.	4620	99
1065	AAB92756	Homo sapiens	Human protein sequence SEQ ID	2856	99

Table 2B  
242

SEQ ID	Hit ID	Species	Description	S score	% Identity
			NO:11216.		
1066	AAE23544	Homo sapiens	Human FAIL protein.	1730	99
1066	AAE23554	Homo sapiens	Human FAIL protein #3.	1728	99
1066	AAE23556	Homo sapiens	Human FAIL protein #5.	1726	99
1067	gi18480772	Mus musculus	olfactory receptor MOR101-2	1267	82
1067	gi18479346	Mus musculus	olfactory receptor MOR101-1	1193	79
1067	AAG72119	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1800.	1043	81
1068	ABB89454	Homo sapiens	Human polypeptide SEQ ID NO 1830.	530	95
1068	ABP41764	Homo sapiens	Human ovarian antigen HE6CR19, SEQ ID NO:2896.	530	95
1068	AAU04352	Homo sapiens	Mammalian toxicological response marker protein #4.	530	95
1069	AAAY33300	Homo sapiens	Human hALK-2 clone HP53 protein.	1572	100
1069	AAR85206	Homo sapiens	Human ALK-2.	1572	100
1069	gi1381584	Bos taurus	activin receptor type I	1572	100
1070	gi16359163	Homo sapiens	Similar to RIKEN cDNA 2310014B08 gene	1332	94
1070	gi18043464	Mus musculus	RIKEN cDNA 2310014B08 gene	1226	77
1070	AAB64401	Homo sapiens	Amino acid sequence of human intracellular signalling molecule INTRA33.	212	35
1071	AAW60043	Homo sapiens	Human MHC class I chain-related gene A (MICA) polypeptide.	1894	93
1071	gi1405893	Homo sapiens	MHC class I chain-related protein A	1894	93
1071	gi16877353	Homo sapiens	MHC class I polypeptide-related sequence A	1838	90
1072	gi15292437	Drosophila melanogaster	LP10272p	439	39
1072	AAB80378	Homo sapiens	Secreted protein encoded by gene #8.	210	28
1072	AAAY87336	Homo sapiens	Human signal peptide containing protein HSPP-113 SEQ ID NO:113.	210	28
1073	AAB58289	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 627.	2303	86
1073	ABB55767	Homo sapiens	Human polypeptide SEQ ID NO 140.	2163	86
1073	AAU39058	Homo sapiens	Human secreted protein pe584_2.	2163	86
1074	AAB58289	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 627.	2303	86
1074	ABB55767	Homo sapiens	Human polypeptide SEQ ID NO 140.	2163	86
1074	AAU39058	Homo sapiens	Human secreted protein pe584_2.	2163	86
1075	AAM93703	Homo sapiens	Human polypeptide, SEQ ID NO: 3632.	1061	93
1075	AAE04780	Homo sapiens	Human vesicle trafficking protein-23 (VETRP-23) protein.	864	100
1075	ABB08160	Homo sapiens	Human cytoskeleton-associated protein (CSAP)-4 (ID: 7472724CD1).	758	87
1076	ABB06255	Homo sapiens	Human G protein-coupled receptor TGR17-6 protein SEQ ID NO:15.	1511	100
1076	ABB06254	Homo sapiens	Human G protein-coupled receptor TGR17-5 protein SEQ ID NO:13.	1511	100
1076	ABB06252	Homo sapiens	Human G protein-coupled receptor TGR17-3 protein SEQ ID NO:7.	1511	100
1077	AAE16786	Homo sapiens	Human transporter and ion channel-	2679	99

Table 2B  
243

SEQ ID	Hit ID	Species	Description	S score	% Identity
			23 (TRICH-23) protein.		
1077	ABB08218	Homo sapiens	Human membrane transporter protein 57256.	1759	68
1077	AAM38697	Homo sapiens	Human polypeptide SEQ ID NO 1842.	1249	53
1078	AAB85029	Homo sapiens	Protein encoded by BAP28 cDNA consisting of exons 1 to 45.	3444	81
1078	AAW54099	Homo sapiens	Homo sapiens BAP28 sequence.	2219	88
1078	AAB92729	Homo sapiens	Human protein sequence SEQ ID NO:11159.	1588	92
1079	ABB07526	Homo sapiens	Human drug metabolizing enzyme (DME) (ID: 1962105CD1).	2580	99
1079	AAM93720	Homo sapiens	Human polypeptide, SEQ ID NO: 3669.	2331	92
1079	gi57806	Rattus sp.	gamma-glutamyltranspeptidase (AA 1-568)	236	28
1080	AAM78536	Homo sapiens	Human protein SEQ ID NO 1198.	6839	100
1080	AAM79520	Homo sapiens	Human protein SEQ ID NO 3166.	6820	99
1080	ABG40303	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 29968.	2092	100
1081	gi20809440	Homo sapiens	sterol O-acyltransferase (acyl-Coenzyme A: cholesterol acyltransferase) 1	1783	65
1081	gi4878022	Homo sapiens	acyl-coenzyme A: cholesterol acyltransferase	1779	65
1081	AAW38416	Homo sapiens	Human acyl-coenzyme A:cholesterol acyltransferase I.	1774	65
1082	gi22002433	Homo sapiens	p150 target of rapamycin (TOR)-scaffold protein containing WD-repeats	7014	100
1082	gi21979456	Homo sapiens	raptor	7014	100
1082	gi22002435	Mus musculus	p150 target of rapamycin (TOR)-scaffold protein containing WD-repeats	6819	96
1083	AAV33301	Homo sapiens	Human hALK-3 clone ONF5 protein.	2647	89
1083	AAR85207	Homo sapiens	Human ALK-3.	2647	89
1083	AAR55368	Homo sapiens	Human Activin receptor-like kinase 3 (hALK-3).	2647	89
1084	AAV33301	Homo sapiens	Human hALK-3 clone ONF5 protein.	1829	92
1084	AAR85207	Homo sapiens	Human ALK-3.	1829	92
1084	AAR55368	Homo sapiens	Human Activin receptor-like kinase 3 (hALK-3).	1829	92
1085	AAW90873	Homo sapiens	Human brain-specific dysferlin protein.	1329	53
1085	AAW90868	Homo sapiens	Human dysferlin protein.	1329	53
1085	AAV82643	Homo sapiens	Human dysferlin protein sequence SEQ ID NO:2.	1329	53
1086	gi19343765	Mus musculus	Similar to dysferlin	1860	47
1086	ABB89615	Homo sapiens	Human polypeptide SEQ ID NO 1991.	1853	47
1086	gi6731235	Homo sapiens	myoferlin	1853	47
1087	AAV92321	Homo sapiens	Human alpha-2-delta-D calcium channel subunit.	5881	99
1087	AAB62262	Homo sapiens	Human calcium channel alpha2delta	5742	99

Table 2B  
244

SEQ ID	Hit ID	Species	Description	S score	% Identity
			subunit.		
1087	AAU01038	Homo sapiens	Human secreted soluble alpha2delta calcium channel subunit #18 protein.	5742	99
1088	gi18676422	Homo sapiens	FLJ00088 protein	3492	96
1088	ABP51380	Homo sapiens	Human MDDT SEQ ID NO 402.	2083	99
1088	gi2104689	Mus musculus	alpha glucosidase II, alpha subunit	1987	53
1089	AAAY01143	Homo sapiens	Secreted protein encoded by gene 9 clone HSIDY06.	238	100
1090	AAAY81261	Homo sapiens	Human DNA structure-specific recognition protein 1 (SSRP1).	3683	100
1090	AAW39212	Homo sapiens	Human SSRP1 protein.	3683	100
1090	AAR38744	Homo sapiens	Human SSRP.	3683	100
1091	gi177814	Homo sapiens	alpha-1-antitrypsin-related protein	1922	90
1091	gi15990507	Homo sapiens	Similar to serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	1408	66
1091	AAAY26925	Homo sapiens	Human alpha1-anti-trypsin type M1 protein.	1407	66
1092	gi16877139	Homo sapiens	Similar to RIKEN cDNA 1300019N10 gene	2273	100
1092	AAB56819	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1397.	1054	100
1092	AAM95577	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 4235.	627	91
1093	gi18605512	Homo sapiens	Similar to CAP-binding protein complex interacting protein 2	1523	100
1093	AAO06814	Homo sapiens	Human polypeptide SEQ ID NO 20706.	1314	100
1093	AAAY57946	Homo sapiens	Human transmembrane protein HTMPN-70.	1128	100
1094	AAM93603	Homo sapiens	Human polypeptide, SEQ ID NO: 3418.	2973	99
1094	gi19571657	Caenorhabditis elegans	similar to Yeast YEH4 like protein	964	41
1094	ABB89291	Homo sapiens	Human polypeptide SEQ ID NO 1667.	590	85
1095	gi4959568	Homo sapiens	nuclear pore complex interacting protein NPIP	1650	94
1095	AAO17206	Homo sapiens	Human secreted protein SEQ ID NO: 105.	1336	79
1095	ABB90262	Homo sapiens	Human polypeptide SEQ ID NO 2638.	872	69
1096	gi18031730	Homo sapiens	GK006	1405	98
1096	AAB92609	Homo sapiens	Human protein sequence SEQ ID NO:10874.	1074	99
1096	AAM40309	Homo sapiens	Human polypeptide SEQ ID NO 3454.	1074	99
1097	AAG03767	Homo sapiens	Human secreted protein, SEQ ID NO: 7848.	612	90
1097	AAB43694	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1139.	612	90
1097	gi285910	Homo sapiens	ATP synthase subunit c precursor	612	90
1098	gi897827	Homo sapiens	iron-responsive element-binding protein/iron regulatory protein 2	4968	99

Table 2B  
245

SEQ ID	Hit ID	Species	Description	S score	% Identity
1098	AAE19851	Homo sapiens	Human wild-type IRP-2 protein.	4909	99
1098	gi897581	Homo sapiens	iron-regulatory protein 2	4909	99
1099	gi3551150	Canine herpesvirus	immediate-early protein	77	28
1099	gi437051	Acipenser transmontanus	vitellogenin	76	23
1099	gi21539886	Arabidopsis thaliana	transcription activator	75	26
1100	AAM84273	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:11866.	328	92
1100	gi3551150	Canine herpesvirus	immediate-early protein	77	28
1100	gi535260	Plasmodium reichenowi	STARP antigen	76	21
1101	gi3551150	Canine herpesvirus	immediate-early protein	77	28
1101	gi437051	Acipenser transmontanus	vitellogenin	76	23
1101	gi535260	Plasmodium reichenowi	STARP antigen	75	21
1102	AAU12296	Homo sapiens	Human PRO7171 polypeptide sequence.	168	67
1102	ABB85001	Homo sapiens	Human PRO28631 protein sequence SEQ ID NO:370.	90	40
1102	ABB95607	Homo sapiens	Human angiogenesis related protein PRO28631 SEQ ID NO: 370.	90	40
1103	gi11558264	Homo sapiens	sphingosine-1-phosphatase	2032	89
1103	gi13447199	Homo sapiens	sphingosine-1-phosphate phosphatase	1994	87
1103	gi15778670	Mus musculus	sphingosine-1-phosphate phosphatase	1721	76
1104	ABB72215	Homo sapiens	Human protein isolated from skin cells SEQ ID NO: 331.	1544	100
1104	ABB72150	Homo sapiens	Human protein isolated from skin cells SEQ ID NO: 189.	1544	100
1104	ABB84843	Homo sapiens	Human PRO301 protein sequence SEQ ID NO:54.	1544	100
1105	AAG03757	Homo sapiens	Human secreted protein, SEQ ID NO: 7838.	506	100
1105	gi178836	Homo sapiens	apolipoprotein C-II	506	100
1105	gi757915	Homo sapiens	apoCII protein	506	100
1106	AAU97773	Homo sapiens	Human Fortilin polypeptide.	794	97
1106	AAR55698	Homo sapiens	Tumor protein p21.	794	97
1106	gi18482460	Sus scrofa	translationally controlled tumor protein	794	97
1107	gi12082725	Mus musculus	B cell phosphoinositide 3-kinase adaptor	3519	84
1107	gi12082723	Gallus gallus	B cell phosphoinositide 3-kinase adaptor	2806	69
1107	gi20987486	Homo sapiens	similar to B cell phosphoinositide 3-kinase adaptor	1829	97
1108	gi21708117	Homo sapiens	similar to hepatocellular carcinoma-associated antigen HCA557b	318	37
1108	gi18252514	Homo sapiens	hepatocellular carcinoma-associated antigen HCA557b	318	37

Table 2B  
246

SEQ ID	Hit ID	Species	Description	S score	% Identity
1108	gi20071260	Mus musculus	RIKEN cDNA 2310038H17 gene	306	37
1109	gi15076511	Homo sapiens	nonmuscle myosin light chain 2	881	99
1109	gi13436446	Homo sapiens	myosin regulatory light chain	881	99
1109	gi22137716	Mus musculus	myosin regulatory light chain	881	99
1110	gi17391357	Homo sapiens	FXYD domain-containing ion transport regulator 7	390	96
1110	gi19354238	Mus musculus	FXYD domain-containing ion transport regulator 1	153	45
1110	gi4206711	Mus musculus	phospholemman precursor	153	45
1111	AAM49040	Homo sapiens	Human testicular development-specific protein 10 (NYD-SP10).	1486	85
1111	gi13272522	Homo sapiens	transcription factor NYD-sp10	1486	85
1111	gi21040409	Homo sapiens	regulatory factor X, 4 (influences HLA class II expression)	1486	85
1112	gi20269720	Homo sapiens	neuropilin and tolloid like-1	1894	99
1112	gi20269724	Mus musculus	neuropilin and tolloid like-1	1839	96
1112	ABB55774	Homo sapiens	Human polypeptide SEQ ID NO 154.	1057	58
1113	AAB61150	Homo sapiens	Human NOV19 protein.	758	98
1113	AAB61149	Homo sapiens	Human NOV18 protein.	758	98
1113	AAAY33297	Homo sapiens	Human membrane spanning protein MSP-4.	758	98
1114	ABB90021	Homo sapiens	Human polypeptide SEQ ID NO 2397.	476	89
1114	ABB11874	Homo sapiens	Human secreted protein homologue, SEQ ID NO:2244.	476	89
1114	AAAY94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	476	89
1115	AAG72407	Homo sapiens	Human OR-like polypeptide query sequence, SEQ ID NO: 2088.	1281	100
1115	AAG72267	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1948.	1281	100
1115	gi21928991	Homo sapiens	seven transmembrane helix receptor	1274	99
1116	AAU80496	Homo sapiens	Human G-coupled receptor (GCRC) protein, Seq ID No 4.	1905	98
1116	ABP51568	Homo sapiens	Human G protein coupled receptor SEQ ID NO:18.	1543	98
1116	AAU85147	Homo sapiens	G-coupled olfactory receptor #8.	1538	100
1117	gi5802817	Homo sapiens	envelope protein	479	77
1117	gi3150438	Human endogenous retrovirus K	pol-env	466	77
1117	gi9558705	Homo sapiens	envelope	466	77
1118	AAE15241	Homo sapiens	Human RNA metabolism protein-4 (RMEP-4).	514	100
1118	ABB89053	Homo sapiens	Human polypeptide SEQ ID NO 1429.	514	100
1118	AAG89341	Homo sapiens	Human secreted protein, SEQ ID NO: 461.	501	99
1119	AAE23979	Homo sapiens	Human LP217 secreted protein.	4563	50
1119	AAB20155	Homo sapiens	Secreted protein SECP1.	4522	50
1119	AAM39295	Homo sapiens	Human polypeptide SEQ ID NO 2440.	4518	50
1120	AAB28199	Homo sapiens	Human HMG-17 non histone chromosomal protein.	429	94

Table 2B  
247

SEQ ID	Hit ID	Species	Description	S score	% Identity
1120	gi32329	Homo sapiens	put. HMG-17 protein	429	94
1120	gi306864	Homo sapiens	high mobility group protein 17	429	94
1121	ABP43105	Homo sapiens	Human ovarian antigen HVCBB19, SEQ ID NO:4237.	456	69
1121	AAE13797	Homo sapiens	Human lung tumour-specific protein SALT-T8.	456	69
1121	AAB44456	Homo sapiens	Human lung tumour-specific antigen encoded by cDNA #71.	456	69
1122	AAM93711	Homo sapiens	Human polypeptide, SEQ ID NO: 3650.	2974	99
1122	ABB89767	Homo sapiens	Human polypeptide SEQ ID NO 2143.	2214	97
1122	gi7303971	Drosophila melanogaster	CG8230-PA	1154	41
1123	AAU76036	Homo sapiens	Human sugar transporter-1 (HST-1) protein sequence.	1055	99
1123	AAB60112	Homo sapiens	Human transport protein TPPT-32.	775	100
1123	AAB61903	Homo sapiens	Atherosclerosis-associated polypeptide.	380	48
1124	AAR28120	Homo sapiens	NKG2 transmembrane protein-D.	725	95
1124	ABB11846	Homo sapiens	Human integral membrane protein homologue, SEQ ID NO:2216.	722	94
1124	gi35063	Homo sapiens	Type II integral membrane protein	722	94
1125	AAM78418	Homo sapiens	Human protein SEQ ID NO 1080.	1878	94
1125	gi21518639	Homo sapiens	TSLC1-like 2	1870	97
1125	gi19068139	Mus musculus	membrane glycoprotein	1849	96
1126	AAB94738	Homo sapiens	Human protein sequence SEQ ID NO:15776.	3079	99
1126	AAM41695	Homo sapiens	Human polypeptide SEQ ID NO 6626.	2456	99
1126	AAM39909	Homo sapiens	Human polypeptide SEQ ID NO 3054.	1272	100
1127	AAB75594	Homo sapiens	Human secreted protein sequence encoded by gene 37 SEQ ID NO:148.	678	99
1127	AAB80437	Homo sapiens	Gene #20 associated peptide #1.	381	98
1127	AAM78175	Homo sapiens	Human bone marrow expressed probe encoded protein SEQ ID NO: 38481.	365	100
1128	gi291529	Mouse cytomegalovirus 1	tegument protein	89	24
1128	gi14573798	Caenorhabditis elegans	C. elegans SRD-60 protein (corresponding sequence C13B7.3)	85	28
1128	gi191992	Mus musculus	APC	79	20
1129	gi20987535	Mus musculus	RIKEN cDNA 3300002C04 gene	741	67
1129	gi20799661	Mus musculus	mucolipin-2	741	67
1129	AAB93412	Homo sapiens	Human protein sequence SEQ ID NO:12616.	632	55
1130	gi19354289	Mus musculus	RIKEN cDNA 2010107G23 gene	111	42
1130	ABB79328	Homo sapiens	Human ovary specific protein SEQ ID NO:125.	107	42
1130	gi15488920	Homo sapiens	Similar to RIKEN cDNA 2010107G23 gene	107	42
1131	AAU70675	Homo sapiens	Human otoferlin.	2432	42

Table 2B  
248

SEQ ID	Hit ID	Species	Description	S score	% Identity
1131	gi4588470	Homo sapiens	otoferlin	2432	42
1131	AAU70673	Homo sapiens	Human otoferlin #2.	2420	42
1132	gi20071179	Homo sapiens	monocyte to macrophage differentiation-associated	1032	77
1132	gi1006665	Homo sapiens	expression associated with monocyte to macrophage differentiation	1028	77
1132	gi18314462	Mus musculus	monocyte to macrophage differentiation-associated	1028	77
1133	AAU95752	Homo sapiens	Human olfactory and pheromone G protein-coupled receptor #239.	1522	92
1133	AAU85278	Homo sapiens	G-coupled olfactory receptor #139.	1499	91
1133	AAU24658	Homo sapiens	Human olfactory receptor AOLFR156.	1499	91
1134	AAE13275	Homo sapiens	Human transporters and ion channels (TRICH)-2.	1472	80
1134	gi17384411	Homo sapiens	ba251O17.3 (similar to aquaporin 7)	1469	80
1134	AAAY70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	1295	73
1135	ABG42409	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 32074.	317	100
1135	AAM32826	Homo sapiens	Peptide #6863 encoded by probe for measuring placental gene expression.	317	100
1135	AAM19638	Homo sapiens	Peptide #6072 encoded by probe for measuring cervical gene expression.	317	100
1136	AAB47977	Homo sapiens	BCY5.	2663	99
1136	gi2463632	Homo sapiens	monocarboxylate transporter homologue MCT6	2574	97
1136	gi21265165	Homo sapiens	solute carrier family 16 (monocarboxylic acid transporters), member 7	602	31
1137	ABB08456	Homo sapiens	Human tumour specific antigenic peptide #2.	705	94
1137	AAU08592	Homo sapiens	Human V-ATPase 16kDa subunit.	705	94
1137	gi14424534	Homo sapiens	ATPase, H <sup>+</sup> transporting, lysosomal (vacuolar proton pump) 16kD	705	94
1138	gi16741167	Mus musculus	RIKEN cDNA 0610010D20 gene	1506	87
1138	gi15080314	Homo sapiens	Similar to RIKEN cDNA 0610010D20 gene	514	100
1138	gi10580053	Halobacterium sp. NRC-1	dihydrodipicolinate synthase; DapA	375	33
1139	AAO14199	Homo sapiens	Human transporter and ion channel TRICH-16.	1425	85
1139	ABB80588	Homo sapiens	Human sbg1020829SGLT protein.	1425	85
1139	AAE06614	Homo sapiens	Human protein having hydrophobic domain, HP03974.	1425	85
1140	ABB90752	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 236.	1890	76
1140	gi6708478	Mus musculus	formin-like protein	1559	65
1140	gi4101720	Mus musculus	lymphocyte specific formin related protein	1532	65
1141	AAB94131	Homo sapiens	Human protein sequence SEQ ID NO:14389.	996	96
1141	gi497984	Bos taurus	Ac45	317	39
1141	AAM93671	Homo sapiens	Human polypeptide, SEQ ID NO:	307	36



Table 2B  
249

SEQ ID	Hit ID	Species	Description	S score	% Identity
			3556.		
1142	AAG67137	Homo sapiens	Amino acid sequence of a human enzyme.	2701	96
1142	gi21707691	Mus musculus	RIKEN cDNA 1500002O20 gene	2645	93
1142	AAB94358	Homo sapiens	Human protein sequence SEQ ID NO:14883.	2579	96
1143	AAW54370	Homo sapiens	G-protein coupled receptor HLTEX11.	1815	100
1143	AAB64854	Homo sapiens	Human secreted protein sequence encoded by gene 36 SEQ ID NO:140.	1792	100
1143	AAB64853	Homo sapiens	Gene 36 human secreted protein homologous amino acid sequence #139.	1792	100
1144	ABB90324	Homo sapiens	Human polypeptide SEQ ID NO 2700.	1321	100
1144	AAU82004	Homo sapiens	Human secreted protein SECP30.	989	78
1144	AAM95005	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 3663.	548	85
1145	AAV13458	Homo sapiens	Amino acid sequence of human Fe65.	3759	100
1145	gi2734083	Homo sapiens	stat-like protein	3759	100
1145	gi3924936	Homo sapiens	Fe65 protein	3759	100
1146	AAB08900	Homo sapiens	Human secreted protein sequence encoded by gene 10 SEQ ID NO:57.	845	90
1146	AAV27071	Homo sapiens	Human JWA protein.	845	90
1146	AAW75110	Homo sapiens	Human secreted protein encoded by gene 54 clone HETGL41.	845	90
1147	AAM93733	Homo sapiens	Human polypeptide, SEQ ID NO: 3697.	764	87
1147	gi2970431	Florometra serratissima	NADH dehydrogenase subunit 4	91	31
1147	gi15128604	Inversidens japanensis	NADH dehydrogenase subunit 4	77	29
1148	AAB93562	Homo sapiens	Human protein sequence SEQ ID NO:12957.	2402	100
1148	gi21626993	Drosophila melanogaster	CG15078-PA	990	39
1148	gi17945442	Drosophila melanogaster	RE18318p	990	39
1149	ABB89832	Homo sapiens	Human polypeptide SEQ ID NO 2208.	1640	99
1149	ABB89833	Homo sapiens	Human polypeptide SEQ ID NO 2209.	838	97
1149	gi16359249	Mus musculus	RIKEN cDNA 1300010M03 gene	630	36
1150	ABB89832	Homo sapiens	Human polypeptide SEQ ID NO 2208.	1640	99
1150	ABB89833	Homo sapiens	Human polypeptide SEQ ID NO 2209.	838	97
1150	gi16359249	Mus musculus	RIKEN cDNA 1300010M03 gene	564	35
1151	ABB78999	Homo sapiens	Human 9-27 protein sequence SEQ ID NO:4483.	569	93
1151	gi1177476	Homo sapiens	interferon-inducible protein	569	93
1151	gi12654159	Homo sapiens	interferon induced transmembrane protein 1 (9-27)	569	93

Table 2B  
250

SEQ ID	Hit ID	Species	Description	S score	% Identity
1152	AAU85153	Homo sapiens	G-coupled olfactory receptor #14.	1730	100
1152	AAU24529	Homo sapiens	Human olfactory receptor AOLFR14.	1730	100
1152	gi21928719	Homo sapiens	seven transmembrane helix receptor	1630	100
1153	gi784997	Homo sapiens	homologue to Drosophila tumour suppressor gene	5300	95
1153	gi1944491	Homo sapiens	homologue of the murine Llg1h gene	4891	90
1153	gi22347352	Rattus norvegicus	RGL1	4878	88
1154	AAM79219	Homo sapiens	Human protein SEQ ID NO 1881.	325	64
1155	AAG63804	Homo sapiens	Amino acid sequence of a human amino acid transporter.	2487	100
1155	gi9309293	Homo sapiens	asc-type amino acid transporter 1	2487	100
1155	gi15277644	Homo sapiens	amino acid transporter	2487	100
1156	gi6760373	Homo sapiens	ODZ3	2323	100
1156	gi4760780	Mus musculus	Ten-m3	2248	96
1156	AAU09891	Homo sapiens	Human heregulin, gamma-HRG.	1202	52
1157	AAAY51559	Homo sapiens	Human RGD1 protein.	724	81
1157	AAW74804	Homo sapiens	Human secreted protein encoded by gene 75 clone HBIAB39.	724	81
1157	gi4877285	Homo sapiens	prenylated Rab acceptor 1 (PRA1)	724	81
1158	gi1780976	Human endogenous retrovirus K	protease	912	59
1158	gi5802824	Homo sapiens	Gag-Pro-Pol protein	909	59
1158	gi5802814	Homo sapiens	Gag-Pro-Pol-Env protein	906	58
1159	AAAY73339	Homo sapiens	HTRM clone 2056042 protein sequence.	867	80
1159	gi13111941	Homo sapiens	vesicle-associated soluble NSF attachment protein receptor (v-SNARE; homolog of S. cerevisiae VTI1)	867	80
1159	gi3861488	Homo sapiens	vesicle soluble NSF attachment protein receptor VTI2	867	80
1160	gi1922891	Mus musculus	alpha 3B chain of laminin-5	10336	75
1160	gi5777581	Homo sapiens	alpha 3B chain of laminin-5	9398	99
1160	AAB48459	Homo sapiens	Human laminin 5 polypeptide, SEQ ID NO: 8.	8690	100
1161	ABP43534	Homo sapiens	Human secreted protein (SCEP) 58.	1984	91
1161	AAG67516	Homo sapiens	Amino acid sequence of a human secreted polypeptide.	1984	91
1161	AAE01167	Homo sapiens	Human gene 4 encoded secreted protein HKAAV61, SEQ ID NO:68.	1984	91
1162	AAM42034	Homo sapiens	Human polypeptide SEQ ID NO 6965.	900	99
1162	AAM40248	Homo sapiens	Human polypeptide SEQ ID NO 3393.	821	100
1162	AAP60254	Homo sapiens	Interferon-pseudo-omega-2.	800	98
1163	AAB19924	Homo sapiens	Human interleukin-1 Hy2 (long version).	818	100
1163	AAB19923	Homo sapiens	Human interleukin-1 Hy2 (extended form, partial sequence).	818	100
1163	AAB19922	Homo sapiens	Human interleukin-1 Hy2 (short version).	818	100
1164	AAB37512	Homo sapiens	Human VEGF C subunit SEQ ID NO: 49.	1061	90

Table 2B  
251

SEQ ID	Hit ID	Species	Description	S score	% Identity
1164	AAAY57035	Homo sapiens	Human A215 amino acid sequence.	1061	90
1164	AAAY94803	Homo sapiens	Human VEGF.	1061	90
1165	ABB84974	Homo sapiens	Human PRO4342 protein sequence SEQ ID NO:316.	770	99
1165	ABB95580	Homo sapiens	Human angiogenesis related protein PRO4342 SEQ ID NO: 316.	770	99
1165	AAB66664	Homo sapiens	Protein encoded by extended B2HFLS20W cDNA library sequence #2.	770	99
1166	gi1321816	Gorilla gorilla	interleukin-8 receptor type B	602	90
1166	AAU10557	Homo sapiens	Human interleukin 8 receptor beta (IL8RB) polypeptide.	599	88
1166	AAU80482	Homo sapiens	Human CXCR2 receptor #1.	599	88
1167	gi1160967	Homo sapiens	palmitoyl-protein thioesterase	1463	92
1167	gi1314355	Homo sapiens	palmitoyl protein thioesterase	1463	92
1167	gi14250054	Homo sapiens	palmitoyl-protein thioesterase 1 (ceroid-lipofuscinosis, neuronal 1, infantile)	1463	92
1168	gi177814	Homo sapiens	alpha-1-antitrypsin-related protein	1952	90
1168	AAAY78890	Homo sapiens	Human alpha1-antitrypsin amino acid sequence.	1450	69
1168	AAW56709	Homo sapiens	Amino acid sequence of the alpha-1-antitrypsin.	1450	69
1169	AAO12931	Homo sapiens	Human polypeptide SEQ ID NO 26823.	346	100
1169	AAO02697	Homo sapiens	Human polypeptide SEQ ID NO 16589.	143	66
1169	AAO08307	Homo sapiens	Human polypeptide SEQ ID NO 22199.	137	80
1170	AAR15222	Homo sapiens	Chronic myelogenous leukaemia-derived myeloid-related protein.	635	100
1170	gi32402	Homo sapiens	HP-1 (AA 1-94)	493	100
1170	gi181529	Homo sapiens	defensin 1	493	100
1171	AAU98513	Homo sapiens	Transmembrane receptor-like protein pTB2185.	71	46
1171	AAU98512	Homo sapiens	Transmembrane receptor-like protein pTB2184.	71	46
1171	gi1001697	Synechocystis sp. PCC 6803	sensory transduction histidine kinase	67	37
1172	AAE14454	Homo sapiens	Human protein phosphatase-4.	2131	99
1172	gi2665458	Mus musculus	protein-tyrosine-phosphatase	1361	63
1172	gi22328117	Homo sapiens	similar to protein-tyrosine-phosphatase homolog DKFZp566K0524.1 - human (fragment)	729	96
1173	AAM25683	Homo sapiens	Human protein sequence SEQ ID NO:1198.	956	96
1173	AAAY48226	Homo sapiens	Human prostate cancer-associated protein 12.	956	96
1173	ABP41699	Homo sapiens	Human ovarian antigen HFTCG52, SEQ ID NO:2831.	950	94
1174	gi20379895	Homo sapiens	Similar to small inducible cytokine A4	124	77
1174	AAM52447	Homo sapiens	HIV_Nef1 fusion protein #14.	122	100
1174	AAM52446	Homo sapiens	HIV_Nef1 fusion protein #13.	122	100

Table 2B  
252

SEQ ID	Hit ID	Species	Description	S score	% Identity
1175	gi20379895	Homo sapiens	Similar to small inducible cytokine A4	305	100
1175	AAM52447	Homo sapiens	HIV_Nef1 fusion protein #14.	123	92
1175	AAM52446	Homo sapiens	HIV_Nef1 fusion protein #13.	123	92
1176	AAG03315	Homo sapiens	Human secreted protein, SEQ ID NO: 7396.	314	100
1176	gi16415877	Octopus salutii	cytochrome oxidase subunit III	70	28
1176	gi602047	Octopus rubescens	cytochrome oxidase subunit III	69	28
1177	AAG03757	Homo sapiens	Human secreted protein, SEQ ID NO: 7838.	453	89
1177	gi178836	Homo sapiens	apolipoprotein C-II	453	89
1177	gi757915	Homo sapiens	apoCII protein	453	89
1178	AAO07986	Homo sapiens	Human polypeptide SEQ ID NO 21878.	73	31
1178	gi4884689	Neisseria meningitidis	lactoferrin-binding protein precursor	70	34
1178	gi172002	Saccharomyces cerevisiae	homologue of bacterial MutS protein	69	24
1179	AAB60502	Homo sapiens	Human cell cycle and proliferation protein CCYPR-50, SEQ ID NO:50.	1656	92
1179	AAB12144	Homo sapiens	Hydrophobic domain protein isolated from WERI-RB cells.	1649	92
1179	gi19880267	Homo sapiens	metallo phosphoesterase	1649	92
1180	AAU91402	Homo sapiens	Human secreted protein sequence #55.	435	100
1180	AAU91376	Homo sapiens	Human secreted protein sequence #29.	435	100
1180	AAU91352	Homo sapiens	Human secreted protein sequence #5.	435	100
1181	AAG01183	Homo sapiens	Human secreted protein, SEQ ID NO: 5264.	275	94
1181	AAO04719	Homo sapiens	Human polypeptide SEQ ID NO 18611.	99	36
1181	AAO03100	Homo sapiens	Human polypeptide SEQ ID NO 16992.	99	42
1182	ABB12063	Homo sapiens	Human secreted protein homologue, SEQ ID NO:2433.	326	100
1182	AAE06730	Homo sapiens	Human CASB765 protein.	200	100
1182	AAU81960	Homo sapiens	Human PRO536.	141	81
1183	AAB75349	Homo sapiens	Human secreted protein #8.	748	98
1183	AAG74502	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5266.	748	98
1183	AAU27661	Homo sapiens	Human protein AFP485790.	748	98
1184	ABB84887	Homo sapiens	Human PRO791 protein sequence SEQ ID NO:142.	1025	90
1184	ABB90305	Homo sapiens	Human polypeptide SEQ ID NO 2681.	1025	90
1184	ABB95493	Homo sapiens	Human angiogenesis related protein PRO791 SEQ ID NO: 142.	1025	90
1185	gi2745733	Homo sapiens	S2P	75	36
1185	gi4164134	Homo sapiens	SP2 metalloprotease	75	36
1185	gi15865970	Varanus spenceri	NADH dehydrogenase subunit 2	73	27
1186	AAU69480	Homo sapiens	Human purified secretory polypeptide #49.	283	92

Table 2B  
253

SEQ ID	Hit ID	Species	Description	S score	% Identity
1186	gi7770223	Homo sapiens	PRO2714	281	84
1187	ABB09144	Homo sapiens	Human phospholipase A2 protein cPLA2 SEQ ID NO:1.	4446	99
1187	ABB07494	Homo sapiens	Human lipid metabolism molecule (LMM) polypeptide (ID: 1281946CD1).	4407	99
1187	AAE05958	Homo sapiens	Human phospholipase-like protein #3.	4215	99
1188	gi14717809	Mus musculus	capsule seleno-protein	96	30
1188	gi14717800	Mus musculus	seleno-protein	96	30
1188	gi2992470	Mus sp.	mitochondrial capsule selenoprotein; MCS	96	30
1189	AAE05095	Homo sapiens	Human inter-alpha trypsin inhibitor (ITI) light chain.	1775	94
1189	AAM79366	Homo sapiens	Human protein SEQ ID NO 3012.	1775	94
1189	AAM78382	Homo sapiens	Human protein SEQ ID NO 1044.	1775	94
1190	gi673422	Homo sapiens	T-cell receptor alpha-chain	1323	93
1190	ABB95403	Homo sapiens	Human P501S specific T cell clone 4E5 Va chain protein SEQ ID NO 906.	1095	77
1190	AAM01298	Homo sapiens	P501S-specific T cell clone 4E5 Va chain T cell receptor amino acid.	1095	77
1191	gi673422	Homo sapiens	T-cell receptor alpha-chain	859	95
1191	gi623119	Macaca mulatta	T-cell receptor alpha	603	86
1191	ABB95403	Homo sapiens	Human P501S specific T cell clone 4E5 Va chain protein SEQ ID NO 906.	593	65
1192	gi53861	Mus musculus	Q300 protein (AA 1-77)	69	41
1193	ABB89241	Homo sapiens	Human polypeptide SEQ ID NO 1617.	247	69
1193	AAB08894	Homo sapiens	Human secreted protein sequence encoded by gene 4 SEQ ID NO:51.	208	57
1193	gi21070180	Danio rerio	envelope protein	102	40
1194	AAG03963	Homo sapiens	Human secreted protein, SEQ ID NO: 8044.	417	80
1194	ABB10412	Homo sapiens	Human cDNA SEQ ID NO: 720.	289	100
1194	ABB10168	Homo sapiens	Human cDNA SEQ ID NO: 476.	289	100
1196	AAM41429	Homo sapiens	Human polypeptide SEQ ID NO 6360.	328	100
1196	AAM39643	Homo sapiens	Human polypeptide SEQ ID NO 2788.	328	100
1196	AAM79610	Homo sapiens	Human protein SEQ ID NO 3256.	328	100
1197	AAE19830	Homo sapiens	Human patched (Ptch) protein.	213	63
1197	AAB67163	Homo sapiens	Human patched protein.	213	63
1197	AAW47157	Homo sapiens	Nevoid basal cell carcinoma syndrome (NBCCS) (PTC) protein.	213	63
1198	AAB25674	Homo sapiens	Human secreted protein sequence encoded by gene 10 SEQ ID NO:63.	643	84
1198	AAB36613	Homo sapiens	Human FLEXHT-35 protein sequence SEQ ID NO:35.	626	80
1198	gi14603247	Homo sapiens	Similar to RIKEN cDNA 5730409G15 gene	626	80
1199	AAO04718	Homo sapiens	Human polypeptide SEQ ID NO 18610.	160	75
1199	AAO04681	Homo sapiens	Human polypeptide SEQ ID NO	160	75

Table 2B  
254

SEQ ID	Hit ID	Species	Description	S score	% Identity
			18573.		
1199	AAO02320	Homo sapiens	Human polypeptide SEQ ID NO 16212.	154	73
1200	AAU12292	Homo sapiens	Human PRO6027 polypeptide sequence.	990	98
1200	AAU27673	Homo sapiens	Human protein AFP235412.	987	99
1200	gi7303340	Drosophila melanogaster	CG4676-PA	189	30
1201	ABG60015	Homo sapiens	Human DITHP polypeptide #73.	281	66
1201	ABG60059	Homo sapiens	Human DITHP polypeptide #117.	249	60
1201	gi21928245	Homo sapiens	seven transmembrane helix receptor	218	60
1202	ABG60015	Homo sapiens	Human DITHP polypeptide #73.	276	68
1202	ABG60059	Homo sapiens	Human DITHP polypeptide #117.	257	47
1202	AAB43928	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1373.	223	55
1203	AAE04368	Homo sapiens	Human kinase (PKIN)-9.	120	85
1203	AAM79153	Homo sapiens	Human protein SEQ ID NO 1815.	120	85
1203	AAV68784	Homo sapiens	Amino acid sequence of a human phosphorylation effector PHSP-16.	120	85
1204	AAU74760	Homo sapiens	Human protease PRTS-20 protein sequence.	1042	98
1204	AAE04733	Homo sapiens	Human protease homologue #1.	1037	98
1204	AAE04734	Homo sapiens	Human protease homologue #2.	1019	97
1205	AAW67842	Homo sapiens	Human secreted protein encoded by gene 36 clone HODCL36.	451	76
1205	gi7717366	Homo sapiens	tryptophan rich protein, congenital heart disease 5 protein CHD5	451	76
1205	gi1946205	Homo sapiens	congenital heart disease 5 protein	445	75
1206	AAG01971	Homo sapiens	Human secreted protein, SEQ ID NO: 6052.	314	100
1206	gi3183989	Lycopersicon esculentum	P69E protein	78	41
1206	gi4200340	Lycopersicon esculentum	P69D protein	77	41
1207	gi14043211	Homo sapiens	Similar to RIKEN cDNA 4931428F04 gene	889	83
1207	gi11276027	Rattus norvegicus	LSC	92	30
1207	gi1389756	Mus musculus	Lsc	92	30
1208	AAV91653	Homo sapiens	Human secreted protein sequence encoded by gene 62 SEQ ID NO:326.	813	79
1208	AAV91512	Homo sapiens	Human secreted protein sequence encoded by gene 62 SEQ ID NO:185.	813	79
1208	AAV44720	Homo sapiens	Human immune system molecule, ISMO-1.	747	78
1209	AAV27648	Homo sapiens	Human secreted protein encoded by gene No. 82.	322	98
1209	gi7959897	Homo sapiens	PRO2379	68	39
1211	AAU19622	Homo sapiens	Human diagnostic and therapeutic polypeptide (DITHP) #208.	339	62
1211	AAB08765	Homo sapiens	A human leukocyte and blood related protein (LBAP).	339	62
1211	AAB74718	Homo sapiens	Human membrane associated protein	314	66

Table 2B  
255

SEQ ID	Hit ID	Species	Description	S score	% Identity
			MEMAP-24.		
1212	AAG65893	Homo sapiens	Amino acid sequence of GSK gene Id 90060.	1901	71
1212	AAB80264	Homo sapiens	Human PRO332 protein.	928	41
1212	AAU12356	Homo sapiens	Human PRO332 polypeptide sequence.	928	41
1213	AAE24240	Homo sapiens	Human 23566 (carboxypeptidase) protein.	1551	99
1213	AAG66547	Homo sapiens	Human secreted metalloproteinase-like polypeptide.	1551	99
1213	AAG66565	Homo sapiens	Human secreted metalloproteinase-like variant polypeptide.	1548	98
1214	AAB82317	Homo sapiens	Human immunoglobulin receptor IRTA4 protein.	528	100
1214	AAB85464	Homo sapiens	Human immunoglobulin domain-containing polypeptide.	528	100
1214	gi15528833	Homo sapiens	Fc receptor-like protein 2	528	100
1215	ABP43492	Homo sapiens	Human secreted protein (SCEP) 16.	1033	100
1215	gi20380668	Homo sapiens	similar to MANNOSE-P-DOLICHOL UTILIZATION DEFECT 1 PROTEIN HOMOLOG	1033	100
1215	AAU27663	Homo sapiens	Human protein AFP285042.	905	90
1216	AAU97104	Homo sapiens	Human MK61 protein, hMK61T4.	979	99
1216	AAU97106	Homo sapiens	Human MK61 protein, hMK61T6.	846	88
1216	AAU97101	Homo sapiens	Human MK61 protein, hMK61T1.	683	83
1217	AAU97106	Homo sapiens	Human MK61 protein, hMK61T6.	833	94
1217	AAU97104	Homo sapiens	Human MK61 protein, hMK61T4.	805	85
1217	AAU97101	Homo sapiens	Human MK61 protein, hMK61T1.	529	92
1218	AAB92697	Homo sapiens	Human protein sequence SEQ ID NO:11091.	71	44
1218	AAB92547	Homo sapiens	Human protein sequence SEQ ID NO:10726.	71	44
1218	gi386638	Mus sp.	mesenchyme fork head 1	70	34
1220	gi1171589	Plasmodium falciparum	frameshift	74	33
1220	gi4512010	Escherichia coli	OrfY	66	50
1221	AAO15424	Homo sapiens	Human genset metabolic gene (GMG-10) protein.	1462	94
1221	AAB50371	Homo sapiens	Human ZACRP7.	1462	94
1221	AAE09444	Homo sapiens	Human SBhACRP30a protein #2.	1462	94
1222	ABG40131	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 29796.	71	34
1222	AAM05935	Homo sapiens	Peptide #4617 encoded by probe for measuring breast gene expression.	71	34
1222	AAM30815	Homo sapiens	Peptide #4852 encoded by probe for measuring placental gene expression.	71	34
1223	gi8850245	Homo sapiens	activated p21cdc42Hs kinase	5605	100
1223	gi2921447	Mus musculus	non-receptor protein tyrosine kinase Ack	5164	92
1223	gi2078388	Bos taurus	Cdc42-associated tyrosine kinase ACK-2	3503	91
1224	AAB84696	Homo sapiens	Amino acid sequence of a human	358	33

Table 2B  
256

SEQ ID	Hit ID	Species	Description	S score	% Identity
			zkun10 polypeptide.		
1224	gi211622	Gallus gallus	alpha-3 collagen type VI	300	32
1224	ABP41277	Homo sapiens	Human ovarian antigen HOCQH66, SEQ ID NO:2409.	281	30
1225	AAB66065	Homo sapiens	Human TANGO 294.	2113	99
1225	AAE11931	Homo sapiens	Human CG162 (or C59) lipase protein #2.	2113	99
1225	AAB66067	Homo sapiens	Human TANGO 294 mature protein.	2015	99
1226	AAM06483	Homo sapiens	Human foetal protein, SEQ ID NO: 214.	283	64
1227	AAR60521	Homo sapiens	Human tetranectin.	869	83
1227	gi37409	Homo sapiens	Tetranectin	869	83
1227	gi825722	Homo sapiens	tetranectin	869	83
1228	gi5790207	Taenia saginata	ATPase subunit 6	68	32
1228	gi8778323	Arabidopsis thaliana	F14J16.24	66	52
1229	AAE01790	Homo sapiens	Human gene 21 encoded secreted protein HDPTW65, SEQ ID NO:111.	142	59
1229	AAE01838	Homo sapiens	Human gene 21 encoded secreted protein HDPTW65, SEQ ID NO:159.	140	57
1229	ABB11479	Homo sapiens	Human reverse transcriptase homologue, SEQ ID NO:1849.	91	55
1230	ABB50466	Homo sapiens	Human secreted protein encoded by gene 166 SEQ ID NO:414.	333	100
1230	AAW88699	Homo sapiens	Secreted protein encoded by gene 166 clone HCEQA68.	333	100
1230	AAU83594	Homo sapiens	Human PRO protein, Seq ID No 6.	327	100
1231	AAG00381	Homo sapiens	Human secreted protein, SEQ ID NO: 4462.	263	91
1231	AAU19357	Homo sapiens	Human G protein-coupled receptor nGPCR-2290.	119	52
1231	AAO13251	Homo sapiens	Human polypeptide SEQ ID NO 27143.	105	39
1232	AAM06558	Homo sapiens	Human foetal protein, SEQ ID NO: 289.	301	98
1233	AAM06562	Homo sapiens	Human foetal protein, SEQ ID NO: 293.	383	100
1233	AAU03519	Homo sapiens	Human protein kinase #19.	76	34
1234	gi13561474	Mertensiella luschni atifi	NADH dehydrogenase subunit 2	70	25
1234	gi13561462	Mertensiella luschni billae	NADH dehydrogenase subunit 2	70	26
1234	gi13561454	Salamandra infraimmaculata	NADH dehydrogenase subunit 2	69	32
1236	gi992917	Glycine max	acetyl CoA carboxylase	67	53
1236	gi1066857	Glycine max	acetyl-CoA carboxylase	65	53
1237	AAE20197	Homo sapiens	Human TIE ligand, FLS139 protein.	1947	94
1237	AAE19828	Homo sapiens	Human FLS 139 ligand protein.	1947	94
1237	ABB80020	Homo sapiens	TIE ligand FLS139 amino acid.	1947	94
1238	AAM06568	Homo sapiens	Human foetal protein, SEQ ID NO: 299.	142	57
1245	gi3873768	Caenorhabditis elegans	contains similarity to Pfam domain: PF01461 (7TM chemoreceptor), Score=28.9, E-value=1.5e-16, N=2	69	28
1245	AAU83682	Homo sapiens	Human PRO protein, Seq ID No 182.	67	39



Table 2B  
257

SEQ ID	Hit ID	Species	Description	S score	% Identity
1245	AAU29165	Homo sapiens	Human PRO polypeptide sequence #142.	67	39
1247	ABB89102	Homo sapiens	Human polypeptide SEQ ID NO 1478.	517	100
1247	AAM38788	Homo sapiens	Human polypeptide SEQ ID NO 1933.	83	21
1247	gi21109777	Xanthomonas axonopodis pv. citri str. 306	transcriptional regulator	83	28
1248	gi15292277	Drosophila melanogaster	LD45324p	70	26
1249	gi298848	Equine herpesvirus 1	glycoprotein C homolog	67	41
1249	gi330808	Equine herpesvirus 1	membrane glycoprotein C	67	41
1249	gi330899	Equine herpesvirus 1	glycoprotein gp13 precursor	67	41
1251	AAO11677	Homo sapiens	Human polypeptide SEQ ID NO 25569.	74	54
1255	gi16551105	Crotalus adamanteus	NADH dehydrogenase subunit 5	66	28
1257	ABB53264	Homo sapiens	Human polypeptide #4.	3044	100
1257	gi18565270	Homo sapiens	Lib	3028	99
1257	gi18565266	Rattus norvegicus	Lib	2551	83
1259	gi335876	Vesicular stomatitis virus	matrix (M) protein	71	26
1259	gi61842	Vesicular stomatitis virus	M protein (aa 1-237)	71	26
1259	gi336034	Vesicular stomatitis virus	M-protein	71	26
1261	AAB95686	Homo sapiens	Human protein sequence SEQ ID NO:18490.	983	99
1261	gi18446901	Drosophila melanogaster	AT07234p	80	32
1262	gi9858058	Tortula ruralis	rehydrin	85	29
1262	gi19882261	Gallus gallus	paranemin	80	27
1262	gi9837383	Bos taurus	retinitis pigmentosa GTPase regulator	76	29
1263	AAU98893	Homo sapiens	Human protease PRTS11.	1600	80
1263	gi965014	Mus musculus	ADAM 4 protein precursor	1297	51
1263	gi1061159	Macaca fascicularis	testicular Metalloprotease-like, Disintegrin-like, Cysteine-rich protein IVa	1253	38
1264	AAE21462	Homo sapiens	Human gene 11 encoded secreted protein HWBBT49, SEQ ID NO:78.	2107	99
1264	AAM80033	Homo sapiens	Human protein SEQ ID NO 3679.	1895	98
1264	AAM79049	Homo sapiens	Human protein SEQ ID NO 1711.	1895	98
1265	AAB44605	Homo sapiens	Human secreted protein sequence encoded by gene 10 SEQ ID NO:70.	93	70
1266	AAO15039	Homo sapiens	Human clusterin protein sequence 1.	1981	87
1266	gi30251	Homo sapiens	SP-40,40 prepropeptide (AA -22 to 427)	1981	87
1266	gi14714741	Homo sapiens	clusterin (complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2,	1981	87

Table 2B  
258

SEQ ID	Hit ID	Species	Description	S score	% Identity
			testosterone-repressed prostate message 2, apolipoprotein J)		
1267	AAB37963	Homo sapiens	Human antithrombin III amino acid sequence.	2136	91
1267	AAAY92224	Homo sapiens	Human antithrombin III.	2136	91
1267	AAR42895	Homo sapiens	Human antithrombin III (wild-type).	2136	91
1268	AAB37963	Homo sapiens	Human antithrombin III amino acid sequence.	2276	96
1268	AAAY92224	Homo sapiens	Human antithrombin III.	2276	96
1268	AAR42895	Homo sapiens	Human antithrombin III (wild-type).	2276	96
1269	gi203519	Rattus norvegicus	cytochrome c oxidase subunit VIc	247	67
1269	gi19354073	Mus musculus	cytochrome c oxidase, subunit VIc	244	67
1269	AAB56523	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1101.	229	61
1270	gi7981304	Homo sapiens	dJ551D2.1.2 (cadherin-like protein VR20, isoform 2)	575	84
1270	AAM96033	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 4691.	512	84
1270	gi9622236	Homo sapiens	cadherin-like protein VR20	272	100
1271	gi6692690	Arabidopsis thaliana	F12K11.14	105	26
1271	gi158696	Drosophila melanogaster	tropomyosin isoform 9D	100	25
1271	gi158695	Drosophila melanogaster	tropomyosin isoform 33 (9C)	100	25
1272	gi6692690	Arabidopsis thaliana	F12K11.14	106	24
1272	gi11935051	Oryctolagus cuniculus	sarcolemmal associated protein 1	99	27
1272	gi1850913	Entamoeba histolytica	myosin heavy chain	98	19
1273	AAM79617	Homo sapiens	Human protein SEQ ID NO 3263.	1481	94
1273	AAM78633	Homo sapiens	Human protein SEQ ID NO 1295.	1481	94
1273	gi16876913	Homo sapiens	mercaptopyruvate sulfurtransferase	1481	94
1274	AAO21899	Homo sapiens	Protein of human protease.	772	49
1274	AAE18722	Homo sapiens	Human DESC1-like serine protease.	767	48
1274	AAB85039	Homo sapiens	Human SER5 protein sequence.	754	47
1275	gi12584839	Homo sapiens	HT036-ISO	1003	94
1275	gi12584841	Homo sapiens	HT036	826	93
1275	gi17985315	Brucella melitensis	HYDROXYPYRUVATE ISOMERASE	512	43
1277	AAB92449	Homo sapiens	Human protein sequence SEQ ID NO:10478.	261	100
1277	AAB83299	Homo sapiens	Human homoglutamine-rich factor 56.	261	100
1277	AAB82342	Homo sapiens	Winged helix/zinc finger transcription factor FOXP1 variant.	261	100
1278	AAM25840	Homo sapiens	Human protein sequence SEQ ID NO:1355.	208	88
1279	AAU72886	Homo sapiens	Human aspartyl protease partial protein sequence #11.	821	98
1279	AAU98884	Homo sapiens	Human protease PRTS2.	821	98
1279	gi7023943	Homo sapiens	down-regulated in gastric cancer	350	35
1280	AAG77975	Homo sapiens	Protein of a human soluble adenylyl	265	37

Table 2B  
259

SEQ ID	Hit ID	Species	Description	S score	% Identity
			cyclase.		
1280	AAE21187	Homo sapiens	Human soluble adenylyl cyclase (SAC) protein.	265	37
1280	AAB81929	Homo sapiens	Human soluble adenylyl cyclase.	265	37
1282	gi4079809	Homo sapiens	HERC2	68	42
1283	ABG45272	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 34937.	153	67
1283	AAM35920	Homo sapiens	Peptide #9957 encoded by probe for measuring placental gene expression.	153	67
1283	AAM20697	Homo sapiens	Peptide #7131 encoded by probe for measuring cervical gene expression.	153	67
1284	AAG81367	Homo sapiens	Human AFP protein sequence SEQ ID NO:252.	849	90
1284	gi18088345	Homo sapiens	Similar to RIKEN cDNA 1110066C01 gene	832	90
1284	gi20381141	Mus musculus	RIKEN cDNA 1110066C01 gene	642	68
1285	AAG81367	Homo sapiens	Human AFP protein sequence SEQ ID NO:252.	939	91
1285	gi18088345	Homo sapiens	Similar to RIKEN cDNA 1110066C01 gene	827	89
1285	gi20381141	Mus musculus	RIKEN cDNA 1110066C01 gene	724	70
1286	AAM79782	Homo sapiens	Human protein SEQ ID NO 3428.	95	36
1286	AAM78798	Homo sapiens	Human protein SEQ ID NO 1460.	93	41
1286	gi177179	Homo sapiens	alpha-2 type VIII collagen	93	41
1287	gi5689766	Homo sapiens	zinc finger 2.2	2092	99
1287	AAM39130	Homo sapiens	Human polypeptide SEQ ID NO 2275.	1048	46
1287	gi3135968	Homo sapiens	b34I8.1 (zinc finger protein 184 (Kruppel-like))	1048	46
1288	ABP43144	Homo sapiens	Human ovarian antigen HVVCD65, SEQ ID NO:4276.	281	86
1288	AAB44228	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1673.	281	86
1288	ABB06807	Homo sapiens	Human nGPCR-Seq1019 protein sequence SEQ ID NO:76.	69	34
1290	AAG03150	Homo sapiens	Human secreted protein, SEQ ID NO: 7231.	307	98
1290	AAW48931	Homo sapiens	Schwannomin-binding protein C-terminal fragment.	286	100
1291	AAB60098	Homo sapiens	Human transport protein TPPT-18.	2360	92
1291	gi1537070	Rattus norvegicus	nucleoporin p54	2292	88
1291	gi7688695	Homo sapiens	nucleoporin p54 protein	2246	89
1292	AAV94621	Homo sapiens	Epidermal growth factor-like variant in skin-2 amino acid sequence.	422	79
1292	ABB72266	Homo sapiens	Human protein isolated from skin cells SEQ ID NO: 417.	415	78
1292	AAB56066	Homo sapiens	Skin cell protein, SEQ ID NO: 417.	415	78
1293	AAM79352	Homo sapiens	Human protein SEQ ID NO 2998.	1018	98
1293	ABB11835	Homo sapiens	Human secreted protein homologue, SEQ ID NO:2205.	1018	98
1293	AAW78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1018	98
1294	AAM99920	Homo sapiens	Human polypeptide SEQ ID NO 36.	669	85

Table 2B  
260

SEQ ID	Hit ID	Species	Description	S score	% Identity
1294	AAM99933	Homo sapiens	Human polypeptide SEQ ID NO 49.	629	82
1294	AAG75526	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6290.	188	38
1295	gi2598167	Homo sapiens	zinc finger protein	2772	99
1295	gi5640019	Mus musculus	zinc finger protein ZFP235	1707	65
1295	gi13277768	Mus musculus	zinc finger protein 93	1434	54
1296	gi4567180	Homo sapiens	BC37295_2 (partial)	2995	100
1296	ABB50238	Homo sapiens	Human transcription factor TRFX-89.	2734	100
1296	gi9502202	Homo sapiens	endothelial zinc finger protein induced by tumor necrosis factor alpha	2734	100
1297	ABG38718	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 28383.	97	55
1297	AAM04653	Homo sapiens	Peptide #3335 encoded by probe for measuring breast gene expression.	97	55
1297	AAM29443	Homo sapiens	Peptide #3480 encoded by probe for measuring placental gene expression.	97	55
1298	AAE14459	Homo sapiens	Human protein phosphatase-9.	601	100
1298	AAM79176	Homo sapiens	Human protein SEQ ID NO 1838.	601	100
1298	AAE14251	Homo sapiens	Human 16051b protein.	563	100
1299	AAB74690	Homo sapiens	Human protease and protease inhibitor PPIM-23.	2420	96
1299	AAB85427	Homo sapiens	Human MPROT45 polypeptide.	2420	96
1299	AAU12243	Homo sapiens	Human PRO4339 polypeptide sequence.	2420	96
1300	AAE16953	Homo sapiens	Human precursor interleukin-18 (Pro-IL-18) protein.	138	92
1300	AAG63830	Homo sapiens	Amino acid sequence of human interleukin 18 (IL-18).	138	92
1300	AAB30541	Homo sapiens	A human IL-18 with a caspase-8 cleavage site.	138	92
1301	AAE05302	Homo sapiens	Human TANGO 457 protein.	901	78
1301	AAE05303	Homo sapiens	Human mature TANGO 457 protein.	889	79
1301	AAE05305	Homo sapiens	Human TANGO 457 protein cytoplasmic domain.	883	78
1302	gi2202	Canis sp.	Clox	68	73
1302	gi6066468	Leishmania major	probable DNA polymerase zeta catalytic component	67	41
1303	gi20068312	Atropa belladonna	maturase	65	34
1304	gi297146	Homo sapiens	retinoic acid receptor gamma 2	274	100
1304	gi18031837	Mesocricetus auratus	retinoic acid receptor gamma-2	256	94
1304	gi200672	Mus musculus	retinoic acid receptor gamma	252	92
1305	AAM39737	Homo sapiens	Human polypeptide SEQ ID NO 2882.	992	99
1305	ABG34053	Homo sapiens	Human Pro peptide #24.	875	100
1305	AAM41523	Homo sapiens	Human polypeptide SEQ ID NO 6454.	875	100
1307	gi212485	Gallus gallus	ovoinhibitor	133	47
1307	ABB72111	Homo sapiens	Human protein isolated from skin cells SEQ ID NO: 150.	132	42
1307	AAB55911	Homo sapiens	Skin cell protein, SEQ ID NO: 150.	132	42

Table 2B  
261

SEQ ID	Hit ID	Species	Description	S score	% Identity
1308	AAU27671	Homo sapiens	Human protein AFP355471.	486	100
1308	gi1256888	Saccharomyces cerevisiae	Ylr086wp	75	30
1308	gi21280444	Homo sapiens	AF15q14	71	27
1309	AAR15222	Homo sapiens	Chronic myelogenous leukaemia-derived myeloid-related protein.	620	100
1309	gi32402	Homo sapiens	HP-1 (AA 1-94)	493	100
1309	gi181529	Homo sapiens	defensin 1	493	100
1310	AAB39114	Homo sapiens	Human secreted protein #22.	347	100
1310	gi2316086	Pisum sativum	gibberellin 3B-hydroxylase	67	43
1310	gi2316018	Pisum sativum	gibberellin 3 beta-hydroxylase	67	43
1311	ABP47852	Homo sapiens	Human polypeptide SEQ ID NO 282.	205	40
1311	ABB90743	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 218.	186	42
1311	gi3127926	Homo sapiens	collagen type VI, alpha 3 chain	186	42
1312	gi16555334	Homo sapiens	Rig protein	1012	100
1312	gi16508176	Homo sapiens	small GTP-binding tumor suppressor 1	1012	100
1312	gi21040535	Homo sapiens	similar to Rig protein	1012	100
1313	gi21741597	Oryza sativa	OSJNBa0052P16.6	84	40
1313	gi21741539	Oryza sativa	OSJNBa0052P16.7	84	40
1313	gi21743203	Oryza sativa	OSJNBa0085C10.10	81	37
1314	gi2689446	Homo sapiens	R27945_1	1889	91
1314	AAM71801	Homo sapiens	Human bone marrow expressed probe encoded protein SEQ ID NO: 32107.	1872	100
1314	AAM79549	Homo sapiens	Human protein SEQ ID NO 3195.	1145	55
1315	AAM79404	Homo sapiens	Human protein SEQ ID NO 3050.	845	79
1315	gi3228237	Homo sapiens	ultra high sulfur keratin	791	69
1315	gi32472	Homo sapiens	high-sulphur keratin	783	76
1316	gi12655446	Homo sapiens	keratin associated protein 4.4	768	78
1316	gi13278909	Homo sapiens	Similar to RIKEN cDNA 1110054P19 gene	738	75
1316	gi12655460	Homo sapiens	keratin associated protein 4.12	738	75
1317	gi12655462	Homo sapiens	keratin associated protein 4.14	1102	88
1317	gi12655452	Homo sapiens	keratin associated protein 4.7	1088	84
1317	gi12655456	Homo sapiens	keratin associated protein 4.9	1002	82
1318	AAM39466	Homo sapiens	Human polypeptide SEQ ID NO 2611.	893	78
1318	AAM41252	Homo sapiens	Human polypeptide SEQ ID NO 6183.	885	78
1318	gi3228237	Homo sapiens	ultra high sulfur keratin	872	73
1319	ABB79480	Homo sapiens	Human zinc finger protein 75.68.	800	65
1319	gi6467200	Homo sapiens	gonadotropin inducible transcription repressor-1	773	63
1319	gi20306351	Homo sapiens	similar to gonadotropin inducible transcription repressor-1	773	63
1320	gi3036963	Ciona savignyi	CsCDC42	163	60
1320	gi21667044	Ustilago maydis	GTP binding protein Cdc42	162	60
1320	gi15072535	Schizophyllum commune	small GTPase CDC42	162	60
1321	AAE02058	Homo sapiens	Human four disulfide core domain (FDCD)-containing protein.	511	43
1321	gi12655452	Homo sapiens	keratin associated protein 4.7	492	43
1321	gi200964	Mus musculus	serine 2 ultra high sulfur protein	481	42

Table 2B  
262

SEQ ID	Hit ID	Species	Description	S score	% Identity
1322	ABB12490	Homo sapiens	Human bone marrow expressed protein SEQ ID NO: 329.	169	72
1323	AAM39144	Homo sapiens	Human polypeptide SEQ ID NO 2289.	378	92
1323	gi5921473	Homo sapiens	G8 protein	378	92
1323	AAM40930	Homo sapiens	Human polypeptide SEQ ID NO 5861.	370	91
1324	gi18446967	Drosophila melanogaster	AT14419p	70	29
1325	gi15277229	Homo sapiens	Homologue to Drosophila photoreceptor protein calphotin	76	34
1325	gi1136400	Homo sapiens	similar to Drosophila photoreceptor cell-specific protein, calphotin.	76	34
1327	AAU07343	Homo sapiens	1-aminocyclopropane carboxylate (ACPC) synthase #12.	262	38
1327	AAU91279	Homo sapiens	Human NOV3a protein.	259	33
1327	gi313720	Ovis aries	KAP5.4 keratin protein	177	26
1328	AAU08325	Homo sapiens	Human granulysin P520 active fragment.	316	75
1328	AAW59874	Homo sapiens	Amino acid sequence of the cDNA clone CAT-1 (HTXET53).	316	75
1328	AAR23732	Homo sapiens	Gene 519 cDNA derived peptide.	316	75
1329	AAM41252	Homo sapiens	Human polypeptide SEQ ID NO 6183.	345	43
1329	AAM39466	Homo sapiens	Human polypeptide SEQ ID NO 2611.	345	43
1329	gi200964	Mus musculus	serine 2 ultra high sulfur protein	342	49
1330	gi13937769	Homo sapiens	Similar to RIKEN cDNA 1200013F24 gene	1256	100
1330	gi7582294	Homo sapiens	BM-011	781	98
1330	AAG67014	Homo sapiens	Human sperm-specific protein EM1, EM6-48.	249	30
1331	gi14718451	Homo sapiens	sialic acid-binding lectin 11	793	70
1331	gi19716086	Homo sapiens	Sialic acid-binding Ig-like lectin Siglec-12	793	70
1331	AAU29082	Homo sapiens	Human PRO polypeptide sequence #59.	551	50
1332	ABP41951	Homo sapiens	Human ovarian antigen HDABR73, SEQ ID NO:3083.	848	98
1332	AAB43821	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1266.	848	98
1332	AAB14201	Homo sapiens	Human placental bikunin protein fragment # 16.	848	98
1333	AAB93164	Homo sapiens	Human protein sequence SEQ ID NO:12091.	2165	98
1333	AAM93693	Homo sapiens	Human polypeptide, SEQ ID NO: 3604.	2159	100
1333	ABB50204	Homo sapiens	Human transcription factor TRFX-55.	1206	57
1334	gi12804907	Homo sapiens	Similar to metaxin 1	1512	100
1334	gi2564913	Homo sapiens	metaxin	1470	90
1334	gi18606009	Mus musculus	metaxin	1333	81
1335	ABB89371	Homo sapiens	Human polypeptide SEQ ID NO 1747.	487	100
1335	ABB90319	Homo sapiens	Human polypeptide SEQ ID NO	381	83

Table 2B  
263

SEQ ID	Hit ID	Species	Description	S score	% Identity
			2695.		
1335	ABP61852	Homo sapiens	Human polypeptide SEQ ID NO 206.	381	83
1336	AAB18456	Homo sapiens	A human TANGO 216 polypeptide clone.	2257	99
1336	AAB18447	Homo sapiens	Amino acid sequence of human TANGO 216 polypeptide.	2257	99
1336	AAB18457	Homo sapiens	A human TANGO 216 polypeptide clone.	2254	99
1337	AAB18456	Homo sapiens	A human TANGO 216 polypeptide clone.	2257	99
1337	AAB18447	Homo sapiens	Amino acid sequence of human TANGO 216 polypeptide.	2257	99
1337	AAB18457	Homo sapiens	A human TANGO 216 polypeptide clone.	2254	99
1338	AAAY86303	Homo sapiens	Human secreted protein HOGCK20, SEQ ID NO:218.	2382	88
1338	AAAY86333	Homo sapiens	Human secreted protein HOGCK20, SEQ ID NO:248.	2215	87
1338	AAB65254	Homo sapiens	Human PRO1379 (UNQ716) protein sequence SEQ ID NO:340.	2117	96
1339	gi20072551	Mus musculus	RIKEN cDNA 4930511J11 gene	412	43
1339	gi12836893	Gallus gallus	IPR328-like protein	150	29
1339	gi17974542	Homo sapiens	voltage-dependent calcium channel gamma-8 subunit	149	25
1340	gi20072551	Mus musculus	RIKEN cDNA 4930511J11 gene	420	45
1340	gi17974542	Homo sapiens	voltage-dependent calcium channel gamma-8 subunit	147	25
1340	gi12836893	Gallus gallus	IPR328-like protein	147	29
1341	AAG89353	Homo sapiens	Human secreted protein, SEQ ID NO: 473.	692	100
1341	ABB11882	Homo sapiens	Human transmembrane protein homologue, SEQ ID NO:2252.	692	100
1341	AAW85737	Homo sapiens	Polypeptide with transmembrane domain.	692	100
1342	ABB12032	Homo sapiens	Human SIGP 2328134 homologue, SEQ ID NO:2402.	1202	82
1342	AAAY21851	Homo sapiens	Human signal peptide-containing protein (SIGP) (clone ID 2328134).	1202	82
1342	gi4101574	Homo sapiens	54TMP	1196	81
1343	gi3002925	Homo sapiens	T cell receptor beta chain	1658	100
1343	AAE13850	Homo sapiens	Human lung tumour-specific N-terminal protein 14F10.	1526	94
1343	AAE13848	Homo sapiens	Human lung tumour-specific T cell receptor beta chain.	1526	94
1344	gi14973269	Streptococcus pneumoniae TIGR4	cell wall surface anchor family protein	413	19
1344	gi560649	Neocallimastix patriciarum	Xylanase B; XYLB	243	19
1344	gi13094677	Mus musculus	ribosome receptor isoform mRRp61	240	22
1345	AAAY07751	Homo sapiens	Human secreted protein fragment encoded from gene 8.	293	100
1345	gi1142588	Trypanosoma brucei	CR3	79	42
1345	gi3037018	Bodo saltans	NADH dehydrogenase subunit 5	78	35

Table 2B  
264

SEQ ID	Hit ID	Species	Description	S score	% Identity
1346	gi18857903	Homo sapiens	TCBA1	867	100
1346	AAG78000	Homo sapiens	Human actin 14.	663	100
1346	ABB89045	Homo sapiens	Human polypeptide SEQ ID NO 1421.	644	98
1347	gi9837433	Homo sapiens	sialic acid binding immunoglobulin-like lectin 8 long splice variant	2206	88
1347	AAW94995	Homo sapiens	SAF-2 polypeptide.	2031	93
1347	gi6980022	Homo sapiens	siglec SAF2	2031	93
1348	gi15451469	Homo sapiens	siglec-like protein splice variant-1	2689	99
1348	gi15217166	Homo sapiens	sialic acid-binding Ig-like lectin 10	2682	99
1348	gi14164613	Homo sapiens	sialic acid binding immunoglobulin-like lectin 10	2356	98
1349	AAU76036	Homo sapiens	Human sugar transporter-1 (HST-1) protein sequence.	1496	89
1349	AAB60112	Homo sapiens	Human transport protein TPPT-32.	775	100
1349	AAB61903	Homo sapiens	Atherosclerosis-associated polypeptide.	445	38
1350	ABB06115	Homo sapiens	Human NS protein sequence SEQ ID NO:207.	357	97
1350	AAV76219	Homo sapiens	Human secreted protein encoded by gene 96.	336	94
1350	gi2906006	Homo sapiens	WASP interacting protein	125	30
1351	AAB08767	Homo sapiens	A human leukocyte and blood related protein (LBAP).	87	27
1351	ABB89384	Homo sapiens	Human polypeptide SEQ ID NO 1760.	86	28
1351	gi576631	Torpedo marmorata	14 kDa transmembrane protein	86	32
1352	AAE16765	Homo sapiens	Human transporter and ion channel-2 (TRICH-2) protein.	312	96
1352	gi7576452	Homo sapiens	potent brain type organic ion transporter	159	37
1352	AAV53009	Homo sapiens	Human secreted protein clone fh149_12 protein sequence SEQ ID NO:24.	153	36
1353	AAU83670	Homo sapiens	Human PRO protein, Seq ID No 158.	2566	99
1353	ABB84896	Homo sapiens	Human PRO1309 protein sequence SEQ ID NO:160.	2566	99
1353	ABB95502	Homo sapiens	Human angiogenesis related protein PRO1309 SEQ ID NO: 160.	2566	99
1354	AAM93665	Homo sapiens	Human polypeptide, SEQ ID NO: 3544.	498	40
1354	AAU29109	Homo sapiens	Human PRO polypeptide sequence #86.	498	40
1354	AAU27785	Homo sapiens	Human full-length polypeptide sequence #110.	498	40
1355	AAB87570	Homo sapiens	Human PRO1268.	603	100
1355	AAU29173	Homo sapiens	Human PRO polypeptide sequence #150.	603	100
1355	AAV78808	Homo sapiens	Hydrophobic domain containing protein clone HP10537 protein sequence.	603	100
1356	gi21518639	Homo sapiens	TSLC1-like 2	1991	97
1356	gi19068139	Mus musculus	membrane glycoprotein	1970	96
1356	AAM78418	Homo sapiens	Human protein SEQ ID NO 1080.	1905	97



Table 2B  
265

SEQ ID	Hit ID	Species	Description	S score	% Identity
1357	AAY49573	Homo sapiens	Human CLA-1 protein sequence.	2503	94
1357	gi397607	Homo sapiens	CLA-1	2503	94
1357	ABB12012	Homo sapiens	Human SR-BI class B scavenger homologue, SEQ ID NO:2382.	2490	94
1358	gi854065	Human herpesvirus 6	U88	340	31
1358	gi21928439	Homo sapiens	seven transmembrane helix receptor	299	35
1358	AAB95124	Homo sapiens	Human protein sequence SEQ ID NO:17122.	277	30
1359	AAE05302	Homo sapiens	Human TANGO 457 protein.	1518	96
1359	AAE05303	Homo sapiens	Human mature TANGO 457 protein.	1394	96
1359	AAE05305	Homo sapiens	Human TANGO 457 protein cytoplasmic domain.	1260	100
1360	gi20799661	Mus musculus	mucolipin-2	2020	76
1360	gi20987535	Mus musculus	RIKEN cDNA 3300002C04 gene	2017	75
1360	gi19072754	Homo sapiens	mucolipin-3	1406	53
1361	AAE04122	Homo sapiens	Human gene 23 encoded secreted protein HE8OK73, SEQ ID NO:108.	210	61
1361	AAE04169	Homo sapiens	Human gene 23 encoded secreted protein HE8OK73, SEQ ID NO:158.	203	60
1361	AAG00392	Homo sapiens	Human secreted protein, SEQ ID NO: 4473.	112	43
1362	AAY27853	Homo sapiens	Human secreted protein encoded by gene No. 101.	274	94
1362	gi21740560	Oryza sativa	OSJNBa0033G16.10	69	47
1362	gi403081	Culex pipiens	reverse transcriptase	66	32
1363	AAM39421	Homo sapiens	Human polypeptide SEQ ID NO 2566.	4775	86
1363	gi17016967	Homo sapiens	NUANCE	4775	86
1363	gi17861384	Homo sapiens	nesprin-2 gamma	4775	86
1364	AAB37381	Homo sapiens	Human secreted protein #10 encoded by cDNA #34.	347	98
1364	AAB95854	Homo sapiens	Human protein sequence SEQ ID NO:18912.	70	37
1365	ABB90157	Homo sapiens	Human polypeptide SEQ ID NO 2533.	136	53
1365	AAW29654	Homo sapiens	Human secreted protein DM406_1.	136	53
1365	gi22328096	Homo sapiens	LOC145053	136	48
1366	AAU76873	Homo sapiens	Human CRF-like protein LP231.	1021	91
1366	ABB53290	Homo sapiens	Human polypeptide #30.	953	65
1366	gi10566471	Mus musculus	Gliacolin	944	65
1367	gi2585988	Homo sapiens	ribonuclease k6 precursor	700	87
1367	gi18088595	Homo sapiens	ribonuclease, RNase A family, k6	700	87
1367	gi2745750	Pan troglodytes	ribonuclease k6 precursor	695	86
1368	AAE09651	Homo sapiens	Human gene 13 encoded lipid metabolism protein HTJN173, SEQ ID NO:45.	484	98
1368	AAG64355	Homo sapiens	Human lambda crystallin.	400	97
1368	ABB89956	Homo sapiens	Human polypeptide SEQ ID NO 2332.	381	98
1369	AAM25241	Homo sapiens	Human protein sequence SEQ ID NO:756.	484	95
1369	AAE09651	Homo sapiens	Human gene 13 encoded lipid metabolism protein HTJN173, SEQ ID NO:45.	352	100

Table 2B  
266

SEQ ID	Hit ID	Species	Description	S score	% Identity
1369	AAG64355	Homo sapiens	Human lambda crystallin.	268	98
1371	gi21685525	Homo sapiens	similar to chloride channel protein SW:CICH_TORMA	4250	94
1371	gi6653659	Oryctolagus cuniculus	chloride channel CLC-6	4155	91
1371	ABB11826	Homo sapiens	Human Cl channel homologue, SEQ ID NO:2196.	4100	91
1372	gi21928599	Homo sapiens	seven transmembrane helix receptor	749	100
1372	AAG71967	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1648.	725	97
1372	AAG71962	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1643.	714	95
1373	ABG45325	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 34990.	278	100
1373	AAM35993	Homo sapiens	Peptide #10030 encoded by probe for measuring placental gene expression.	278	100
1373	AAM20725	Homo sapiens	Peptide #7159 encoded by probe for measuring cervical gene expression.	278	100
1374	AAU12071	Homo sapiens	Human PHT1 variant protein from Caco-2 cells.	920	100
1374	AAU12070	Homo sapiens	Human PHT1 variant protein from BeWo cells.	920	100
1374	AAU12069	Homo sapiens	Human PHT1 protein splice variant.	920	100
1375	ABB77396	Homo sapiens	Human cathepsin L.	1597	87
1375	AAW47031	Homo sapiens	Human procathepsin L.	1597	87
1375	gi29715	Homo sapiens	pro-(cathepsin L)	1597	87
1376	AAU11764	Homo sapiens	Human alpha1a adrenergic receptor-like GPCR.	2554	98
1376	AAG64126	Homo sapiens	Human G protein-coupled receptor GPRv72.	2554	98
1376	AAU04369	Homo sapiens	Human G-protein coupled receptor, hRUP15.	2554	98
1377	gi18676524	Homo sapiens	FLJ00159 protein	164	52
1377	gi21392066	Drosophila melanogaster	RE04357p	139	34
1377	AAB94312	Homo sapiens	Human protein sequence SEQ ID NO:14783.	82	38
1378	ABB08011	Homo sapiens	Human secretin receptor-like GPCR.	2759	100
1378	AAU79494	Homo sapiens	Human G protein-coupled receptor TGR18-3.	2582	97
1378	AAE18652	Homo sapiens	Human G-protein coupled receptor (GCREC-13).	2569	96
1379	AAY30735	Homo sapiens	Amino acid sequence of a human secreted protein.	280	100
1379	AAY40092	Homo sapiens	Peptide sequence derived from a human secreted protein.	65	100
1379	gi17511709	Homo sapiens	down-regulated by Ctnnb1, a	65	26
1380	gi2459682	Homo sapiens	MAGE-B4	883	54
1380	gi3687196	Homo sapiens	MAGE-B1	864	54
1380	gi4033512	Homo sapiens	DAM10=DSS-AHC critical interval MAGE superfamily protein	863	54
1381	AAE21804	Homo sapiens	Human TREK2 protein.	2791	100
1381	gi19716290	Homo sapiens	potassium channel TREK2 splice variant b	2784	99

Table 2B

267

SEQ ID	Hit ID	Species	Description	S score	% Identity
1381	AAO14193	Homo sapiens	Human transporter and ion channel TRICH-10.	2747	98
1382	ABB11297	Homo sapiens	Human Coxsackie adenovirus receptor homologue, SEQ ID NO:1667.	699	97
1382	gi14279421	Danio rerio	coxsackievirus and adenovirus receptor-like protein	366	32
1382	ABB84956	Homo sapiens	Human PRO5723 protein sequence SEQ ID NO:280.	355	33
1383	AAU99321	Homo sapiens	Human protooncogene KG-19 protein.	753	100
1383	gi18466808	Homo sapiens	cervical cancer 1 proto-oncogene-binding protein KG19	753	100
1383	gi21961229	Homo sapiens	BRI3 binding protein	753	100
1384	AAM06866	Homo sapiens	Human foetal protein, SEQ ID NO: 1074.	1131	98
1384	gi15099951	Mus musculus	diacylglycerol acyltransferase 2	953	51
1384	ABB75677	Homo sapiens	Breast protein-eukaryotic conserved gene 1 (BSTP-ECG1) protein.	945	50
1385	AAG68335	Homo sapiens	Human CSP2 protein SEQ ID NO:4.	1354	99
1385	gi19525540	Homo sapiens	lymphocyte effector toxicity activation ligand	1330	98
1385	AAAY36071	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 456.	1252	92
1386	AAAY94527	Homo sapiens	Human statherin protein.	260	83
1386	gi338611	Homo sapiens	statherin precursor	260	83
1386	gi338508	Homo sapiens	statherin protein	260	83
1387	AAB94721	Homo sapiens	Human protein sequence SEQ ID NO:15739.	1011	100
1387	gi17384256	Homo sapiens	mucin 5	103	28
1387	gi1334899	Human herpesvirus 4	BDLF3 late reading frame 9xNXT/S	100	25
1388	ABP41842	Homo sapiens	Human ovarian antigen HTFML39, SEQ ID NO:2974.	918	98
1388	AAB73302	Homo sapiens	Human cyclophilin B (CypB) C-terminal deletion mutant, CypB-AIAKE.	918	98
1388	AAB73301	Homo sapiens	Human cyclophilin B (CypB).	918	98
1389	AAE07112	Homo sapiens	Human gene 6 encoded secreted protein fragment, SEQ ID NO:129.	2470	99
1389	AAM93449	Homo sapiens	Human polypeptide, SEQ ID NO: 3098.	2378	99
1389	AAM93823	Homo sapiens	Human polypeptide, SEQ ID NO: 3881.	2374	99
1390	gi16589056	Homo sapiens	type II gonadotropin-releasing hormone receptor	1376	94
1390	gi19697896	Homo sapiens	GnRH receptor II 5TM	1365	93
1390	AAU10819	Homo sapiens	Human Type II GnRH-R splice variant 1 protein #2.	1363	93
1391	gi16359249	Mus musculus	RIKEN cDNA 1300010M03 gene	2226	91
1391	AAM93450	Homo sapiens	Human polypeptide, SEQ ID NO: 3100.	587	34
1391	ABB89832	Homo sapiens	Human polypeptide SEQ ID NO 2208.	522	39
1392	AAE04896	Homo sapiens	Human transporter and ion channel-9	825	100

Table 2B

268

SEQ ID	Hit ID	Species	Description	S score	% Identity
			(TRICH-9) protein.		
1392	AAE04899	Homo sapiens	Human transporter and ion channel-12 (TRICH-12) protein.	688	52
1392	AAE06616	Homo sapiens	Human protein having hydrophobic domain, HP10735.	688	52
1393	gi21928487	Homo sapiens	seven transmembrane helix receptor	1558	100
1393	AAG72603	Homo sapiens	Human OR-like polypeptide query sequence, SEQ ID NO: 2284.	1433	89
1393	AAG71515	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1196.	1433	89
1394	ABB89241	Homo sapiens	Human polypeptide SEQ ID NO 1617.	195	57
1394	AAB08894	Homo sapiens	Human secreted protein sequence encoded by gene 4 SEQ ID NO:51.	165	59
1394	gi21070180	Danio rerio	envelope protein	111	48
1395	gi15157307	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_C_3928p	85	21
1395	gi9295309	frog adenovirus 1	pIIIa protein	82	41
1395	ABB08204	Homo sapiens	Human lipid metabolism enzyme-4 (LME-4).	76	27
1397	AAG65914	Homo sapiens	Amino acid sequence of GSK gene Id 27142.	5760	97
1397	gi19570398	Homo sapiens	hDDM36	5760	97
1397	AAU77405	Homo sapiens	Human NOV1 protein, homologue of NOPE/PUNC Ig proteins.	5619	95
1398	gi15292481	Drosophila melanogaster	SD03655p	1165	59
1398	AAB88372	Homo sapiens	Human membrane or secretory protein clone PSEC0108.	881	67
1399	ABB07553	Homo sapiens	BSTP-CAD fragment determined from I.M.A.G.E clone 52071.	91	29
1399	ABB07552	Homo sapiens	Human BSTP-CAD polypeptide.	91	29
1399	gi21959343	Yersinia pestis KIM	basal-body MS (membrane and supramembrane)-ring and collar protein	69	35
1401	AAB95124	Homo sapiens	Human protein sequence SEQ ID NO:17122.	148	34
1401	gi854065	Human herpesvirus 6	U88	139	40
1401	AAO09309	Homo sapiens	Human polypeptide SEQ ID NO 23201.	137	26
1402	AAE01332	Homo sapiens	Human gene 18 encoded secreted protein fragment, SEQ ID NO:197.	222	100
1402	AAE01299	Homo sapiens	Human gene 18 encoded secreted protein HFIIN69, SEQ ID NO:162.	222	100
1402	AAE01249	Homo sapiens	Human gene 18 encoded secreted protein HFIIN69, SEQ ID NO:111.	222	100
1403	AAM06589	Homo sapiens	Human foetal protein, SEQ ID NO: 320.	237	100
1404	AAU10510	Homo sapiens	Human leukocyte immunoglobulin receptor-like (LIR-like) protein #8.	1411	100
1404	AAU10511	Homo sapiens	Human leukocyte immunoglobulin receptor-like (LIR-like) protein #11.	595	89

Table 2B

269

SEQ ID	Hit ID	Species	Description	S score	% Identity
1404	AAO09486	Homo sapiens	Human polypeptide SEQ ID NO 23378.	486	98
1406	gi20380511	Mus musculus	RIKEN cDNA 3110007F17 gene	293	33
1406	gi13377867	Gallus gallus	claudin-3	107	24
1406	gi15553375	Danio rerio	claudin h	106	23
1407	ABB90122	Homo sapiens	Human polypeptide SEQ ID NO 2498.	1029	84
1407	AAY32204	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2132179.	1029	84
1407	AAY78801	Homo sapiens	Hydrophobic domain containing protein clone HP00631 amino acid sequence.	1029	84
1408	gi13543940	Homo sapiens	Similar to RIKEN cDNA 2610017G09 gene	2229	96
1408	AAB12138	Homo sapiens	Hydrophobic domain protein isolated from HT-1080 cells.	2218	96
1408	AAB88466	Homo sapiens	Human membrane or secretory protein clone PSEC0260.	2212	96
1409	AAM38643	Homo sapiens	Human polypeptide SEQ ID NO 1788.	3168	99
1409	AAU81961	Homo sapiens	Human PRO943.	2441	98
1409	AAU77790	Homo sapiens	Human PRO943 protein.	2441	98
1410	AAU78083	Homo sapiens	Human interleukin 22 receptor (IL-22R) protein sequence.	1699	100
1410	AAU76905	Homo sapiens	Human Z-cytos-II protein.	1699	100
1410	AAB87607	Homo sapiens	Human PRO20233.	1699	100
1411	AAO20532	Homo sapiens	Protein of the human TFM-2 gene sequence.	777	100
1411	AAE21184	Homo sapiens	Human TRICH-28 protein.	777	100
1411	gi18640047	Homo sapiens	aromatic amino acid transporter	777	100
1412	AAE13280	Homo sapiens	Human transporters and ion channels (TRICH)-7.	1317	78
1412	AAB47271	Homo sapiens	hOAT1.	627	41
1412	AAY44278	Homo sapiens	Human organic anion transporter.	627	41
1413	AAM79977	Homo sapiens	Human protein SEQ ID NO 3623.	100	31
1413	gi2764507	Locusta migratoria	nicotinic acetylcholine receptor, alpha1 subunit	100	36
1413	gi1766077	Gallus gallus	winged helix protein CWH-3	100	37
1414	ABP41590	Homo sapiens	Human ovarian antigen HAZAR95, SEQ ID NO:2722.	1522	100
1414	AAB43682	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1127.	1522	100
1414	gi5931730	Homo sapiens	phosphate carrier	1522	100
1415	gi1764015	Ciona intestinalis	COS41.5	314	42
1415	AAM42167	Homo sapiens	Human polypeptide SEQ ID NO 7098.	272	30
1415	AAE03484	Homo sapiens	Human gene 12 encoded secreted protein HETHW90, SEQ ID NO:167.	250	29
1416	AAB82315	Homo sapiens	Human immunoglobulin receptor isoform IRTA2c.	1130	88
1416	gi15277746	Homo sapiens	Fc receptor-like protein 5	1130	88
1416	gi13591714	Homo sapiens	immunoglobulin superfamily receptor translocation associated protein 2c	1130	88

Table 2B

270

SEQ ID	Hit ID	Species	Description	S score	% Identity
1417	AAAY27182	Homo sapiens	Human RNF5 polypeptide.	877	90
1417	AAW69601	Homo sapiens	Human zinc binding protein ZB-2.	877	90
1417	gi13366064	Homo sapiens	HsRma1	877	90
1418	AAE17493	Homo sapiens	Human secretion and trafficking protein-2 (SAT-2).	1446	100
1418	ABG60227	Homo sapiens	Human Mitsugumin29-like protein NOV3a.	1384	97
1418	AAU96848	Homo sapiens	Human NOV3a protein variant.	1378	97
1419	gi13452508	Mus musculus	claudin 14	371	40
1419	AAU77764	Homo sapiens	Tumour associated antigenic target polypeptide (TAT) 155.	370	40
1419	AAU29200	Homo sapiens	Human PRO polypeptide sequence #177.	370	40
1420	ABG34068	Homo sapiens	Human Pro peptide #39.	697	87
1420	gi21205866	Homo sapiens	T-cell activation leucine repeat-rich protein; TA-LRRP	697	87
1420	AAE17132	Homo sapiens	Human adenylyl and guanylyl cyclase (ADGUC)-4.	505	84
1421	AAU78091	Homo sapiens	Human nonerythroid Rh glycoprotein RhBG protein sequence.	1891	88
1421	gi15718471	Homo sapiens	Rh type B glycoprotein	1891	88
1421	gi14346006	Pan troglodytes	Rh type B glycoprotein	1886	88
1422	AAM00949	Homo sapiens	Human bone marrow protein, SEQ ID NO: 425.	215	46
1422	gi14209836	Mus musculus	ATP-binding cassette transporter sub-family A member 7	175	40
1422	AAO14210	Homo sapiens	Human transporter and ion channel TRICH-27.	174	40
1423	ABB84987	Homo sapiens	Human PRO6006 protein sequence SEQ ID NO:342.	718	99
1423	ABG34059	Homo sapiens	Human Pro peptide #30.	718	99
1423	ABB95593	Homo sapiens	Human angiogenesis related protein PRO6006 SEQ ID NO: 342.	718	99
1424	gi18539467	Homo sapiens	nucleolar RNA-associated protein beta	3851	98
1424	gi18539465	Homo sapiens	nucleolar RNA-associated protein alpha	3851	98
1424	gi20988377	Homo sapiens	nucleolar RNA-associated protein	3823	98
1425	AAE15635	Homo sapiens	Human G-protein coupled receptor-5 (GCREC-5) protein.	499	92
1425	AAB66272	Homo sapiens	Human TANGO 378 SEQ ID NO: 29.	499	92
1425	AAE03406	Homo sapiens	Human secreted protein fragment, SEQ ID NO:64.	499	92
1426	gi18676704	Homo sapiens	FLJ00251 protein	4114	99
1426	gi433383	Tripneustes gratilla	dynein heavy chain isotype 5A	406	24
1426	gi6706264	Leishmania major	dynein heavy chain	405	23
1427	AAE23980	Homo sapiens	Human LP220 secreted protein.	2964	99
1427	AAB70072	Homo sapiens	Human secreted protein #11.	2475	99
1427	AAE17484	Homo sapiens	Human leucine-rich repeat-8 (ZLRR8) protein #2.	2407	100
1428	ABB97996	Homo sapiens	Human severing stimulating factor Kda subunit 86.35.	415	39

Table 2B

271

SEQ ID	Hit ID	Species	Description	S score	% Identity
1428	gi21618514	Homo sapiens	Similar to oxidation resistance 1	415	39
1428	gi13540300	Mus musculus	nucleolar protein C7B	408	34
1429	ABB90382	Homo sapiens	Human polypeptide SEQ ID NO 2758.	352	63
1429	ABP41785	Homo sapiens	Human ovarian antigen HCFCS40, SEQ ID NO:2917.	352	63
1429	AAM40391	Homo sapiens	Human polypeptide SEQ ID NO 3536.	131	32
1430	AAY07751	Homo sapiens	Human secreted protein fragment encoded from gene 8.	293	100
1430	gi1142588	Trypanosoma brucei	CR3	78	43
1430	gi3037018	Bodo saltans	NADH dehydrogenase subunit 5	76	33
1431	ABB04471	Homo sapiens	Human endoprotease 31.	940	100
1431	AAM93525	Homo sapiens	Human polypeptide, SEQ ID NO: 3259.	243	36
1431	ABB89344	Homo sapiens	Human polypeptide SEQ ID NO 1720.	219	29
1432	AAB88388	Homo sapiens	Human membrane or secretory protein clone PSEC0131.	306	43
1432	gi20810493	Homo sapiens	Similar to RIKEN cDNA 2810417M05 gene	306	43
1432	AAE21272	Homo sapiens	Human gene 16 encoded secreted protein fragment, SEQ ID NO:138.	207	55
1433	AAB88388	Homo sapiens	Human membrane or secretory protein clone PSEC0131.	308	42
1433	gi20810493	Homo sapiens	Similar to RIKEN cDNA 2810417M05 gene	308	42
1433	AAE21272	Homo sapiens	Human gene 16 encoded secreted protein fragment, SEQ ID NO:138.	209	54
1434	ABP43139	Homo sapiens	Human ovarian antigen HVVBT60, SEQ ID NO:4271.	235	77
1434	ABP42815	Homo sapiens	Human ovarian antigen HPCOO95, SEQ ID NO:3947.	235	77
1434	ABP42409	Homo sapiens	Human ovarian antigen HOCOC38, SEQ ID NO:3541.	235	77
1435	gi19171162	Homo sapiens	ventrroid transmembrane protein	1873	99
1435	AAE14571	Homo sapiens	Human rhomboid related protein, RRP3.	1869	99
1435	gi19171160	Mus musculus	ventrroid transmembrane protein	1800	94
1436	gi7106866	Homo sapiens	HSPC238	237	56
1436	AAM92052	Homo sapiens	Human digestive system antigen SEQ ID NO: 1401.	235	72
1436	gi13542707	Mus musculus	RIKEN cDNA 2500002L14 gene	195	47
1437	AAE18212	Homo sapiens	Human MOL4 protein.	6140	99
1437	AAG68293	Homo sapiens	Human semaphorin G-like NHP protein SEQ ID NO:10.	6134	99
1437	AAG68294	Homo sapiens	Human semaphorin G-like NHP protein SEQ ID NO:12.	6034	98
1438	AAU95381	Homo sapiens	Human calcium transport protein CaTrF2E11.	4607	99
1438	AAY96479	Homo sapiens	Human vanilloid receptor 3.	4607	99
1438	gi21690507	unidentified	CaTrF2E11	4607	99
1439	gi11055322	Homo sapiens	vanilloid receptor-related osmotically activated channel	3324	100

Table 2B

272

SEQ ID	Hit ID	Species	Description	S score	% Identity
1439	AAU74935	Homo sapiens	Amino acid sequence of human vanilloid receptor-like protein 2a (VRL-2a).	3318	99
1439	ABB79191	Homo sapiens	Human VR4 protein SEQ ID NO:2.	3318	99
1440	gi14574118	Caenorhabditis elegans	C. elegans DPY-19 protein (corresponding sequence F22B7.10)	76	24
1440	AAU83621	Homo sapiens	Human PRO protein, Seq ID No 60.	74	24
1440	AAO05826	Homo sapiens	Human polypeptide SEQ ID NO 19718.	70	26
1441	gi4235228	Mus musculus	leucine zipper-EF-hand containing transmembrane protein 1	497	59
1441	gi15680275	Homo sapiens	Similar to leucine zipper-EF-hand containing transmembrane protein 1	494	60
1441	gi18204589	Homo sapiens	leucine zipper-EF-hand containing transmembrane protein 1	494	60
1442	AAE17482	Homo sapiens	Human leucine-rich repeat-7 (ZLRR7) protein.	1107	100
1442	ABB11242	Homo sapiens	Human SLIT-2 homologue, SEQ ID NO:1612.	653	99
1442	AAB07469	Homo sapiens	A human leucine-rich repeat protein designated Zlrr3.	443	36
1443	AAU97218	Homo sapiens	Human G protein-coupled receptor, TGR25.	1645	87
1443	AAE23415	Homo sapiens	Human G-protein coupled receptor-7 (GCREC-7).	1645	87
1444	gi15487341	Escherichia coli	macrolide-specific ABC-type efflux carrier	697	91
1444	gi4062463	Escherichia coli	ABC transporter probable ATP-binding subunit homolog	697	91
1445	gi17978985	Arabidopsis thaliana	At1g43690/F2J6_4	75	22
1445	gi4239789	Treponema maltophilum	major sheath protein	74	23
1446	gi21426922	Homo sapiens	PELP1	4113	95
1446	AAW31185	Homo sapiens	Human p160 polypeptide 160.1.	3176	87
1446	gi3168604	Homo sapiens	proline and glutamic acid rich nuclear protein isoform	1641	99
1447	AAB94495	Homo sapiens	Human protein sequence SEQ ID NO:15188.	68	36
1447	AAG67254	Homo sapiens	Amino acid sequence of a human liver-associated gene.	68	36
1447	gi21489412	Bombus hypocrita sapporoensis	cytochrome oxidase 1	66	35
1448	AAB80418	Homo sapiens	Gene #4 associated peptide #1.	1998	100
1448	AAB65162	Homo sapiens	Human PRO290 (UNQ253) protein sequence SEQ ID NO:33.	1972	100
1448	AAV66639	Homo sapiens	Membrane-bound protein PRO290.	1972	100
1449	gi16117372	Macropodid herpesvirus 1	ICP4	95	27
1449	gi4063766	Aspergillus nidulans	chitinase	92	24
1449	gi15277229	Homo sapiens	Homologue to Drosophila photoreceptor protein calphotin	89	31
1450	ABB09760	Homo sapiens	Amino acid sequence of human	1029	100



Table 2B  
273

SEQ ID	Hit ID	Species	Description	S score	% Identity
			hearing defect related protein 38.39.		
1450	AAU27660	Homo sapiens	Human protein AFP671052.	889	100
1450	AAG74151	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:4915.	748	100
1452	AAE21175	Homo sapiens	Human TRICH-19 protein.	1041	93
1452	gi15559050	Ethmostigmus rubripes	ADP-ATP translocator	733	68
1452	gi7542476	Xenopus laevis	adenine nucleotide translocase	702	67
1454	gi4929597	Homo sapiens	CGI-64 protein	1879	96
1454	gi6995987	Homo sapiens	mitochondrial carrier homolog 1 isoform a	1818	99
1454	gi6995989	Homo sapiens	mitochondrial carrier homolog 1 isoform b	1790	95
1455	gi17225210	Podospira anserina	beta transducin-like protein HET-D2Y	247	25
1455	gi886024	Thermomonospora curvata	PkwA	239	29
1455	gi17131893	Nostoc sp. PCC 7120	WD-repeat protein	238	26
1456	AAB36840	Homo sapiens	Human insulin receptor-related receptor protein with signal peptide.	6736	98
1456	ABB11702	Homo sapiens	Human insulin receptor-related protein homologue, SEQ ID NO:2072.	6731	99
1456	AAB36836	Homo sapiens	Human insulin receptor-related receptor protein.	6731	99
1457	ABB11803	Homo sapiens	Human GPI-122 homologue, SEQ ID NO:2173.	6420	99
1457	AAY50125	Homo sapiens	Human glycoposphatidylinositol-anchored protein GPI-122.	6323	100
1457	AAB94751	Homo sapiens	Human protein sequence SEQ ID NO:15805.	5872	99
1458	AAU00023	Homo sapiens	Human activated T-lymphocyte associated sequence 2, ATLAS-2.	3623	99
1458	AAE04546	Homo sapiens	Human G-protein coupled receptor-2 (GCREC-2) protein.	3325	89
1458	gi19387136	Homo sapiens	PYRIN-containing APAF1-like protein 5	3316	89
1459	AAE14719	Homo sapiens	Human carbohydrate-associated protein (CARBAP)-1.	3132	99
1459	AAB92508	Homo sapiens	Human protein sequence SEQ ID NO:10631.	2742	100
1459	AAB95542	Homo sapiens	Human protein sequence SEQ ID NO:18155.	2687	99
1460	gi2960194	Human T-cell lymphotropic virus type 2b	rex	81	24
1460	gi404041	Human T-lymphotropic virus 2	rex protein	81	24
1460	gi348140	Human T-lymphotropic virus 2	rex	81	24
1461	gi1103953	Homo sapiens	MHC class I HLA-B*1523	71	28
1461	gi1815616	Homo sapiens	HLA-B-1521	71	28

Table 2B  
274

SEQ ID	Hit ID	Species	Description	S score	% Identity
1461	gi17467060	Homo sapiens	MHC class I	71	28
1462	gi11066090	Homo sapiens	matrix metalloprotease MMP-27	1188	85
1462	AAU00469	Homo sapiens	Human TANGO 210 protein.	1185	85
1462	AAU12267	Homo sapiens	Human PRO5992 polypeptide sequence.	1185	85
1463	gi601948	Drosophila melanogaster	Inscuteable	132	23
1463	gi7291194	Drosophila melanogaster	CG11312-PA	132	23
1463	gi1657962	Drosophila melanogaster	Nem	123	24
1464	AAB45378	Homo sapiens	Human secreted protein sequence encoded by gene 38 SEQ ID NO:130.	433	96
1464	gi406058	Mus musculus	protein kinase	216	53
1464	AAE16277	Homo sapiens	Human kinase PKIN-23 protein.	213	53
1465	AAG00866	Homo sapiens	Human secreted protein, SEQ ID NO: 4947.	123	54
1466	AAV97293	Homo sapiens	Lipid associated protein (LIPAP) 3335404CD1.	2178	68
1466	AAM39997	Homo sapiens	Human polypeptide SEQ ID NO 3142.	1214	43
1466	AAB24231	Homo sapiens	Human vesicle associated protein 10 SEQ ID NO:10.	1214	43
1467	ABP41928	Homo sapiens	Human ovarian antigen H6EEO05, SEQ ID NO:3060.	922	96
1467	AAB94535	Homo sapiens	Human protein sequence SEQ ID NO:15273.	922	96
1467	AAM79530	Homo sapiens	Human protein SEQ ID NO 3176.	922	96
1468	AAB94336	Homo sapiens	Human protein sequence SEQ ID NO:14836.	694	77
1468	AAB97250	Homo sapiens	HOMO RNA cyclase 41 protein.	694	77
1468	gi12654401	Homo sapiens	Similar to RNA cyclase homolog	694	77
1469	ABB05645	Homo sapiens	Human thyroglobulin 38 protein SEQ ID NO:2.	273	98
1469	AAM89857	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:17450.	81	43
1470	AAO04234	Homo sapiens	Human polypeptide SEQ ID NO 18126.	185	76
1470	AAM92673	Homo sapiens	Human digestive system antigen SEQ ID NO: 2022.	184	97
1470	gi65265	Xenopus laevis	a xenopus upstream binding factor	130	30
1471	ABB12219	Homo sapiens	Human secreted protein homologue, SEQ ID NO:2589.	182	97
1471	ABG60080	Homo sapiens	Human DITHP polypeptide #138.	153	82
1471	ABB90307	Homo sapiens	Human polypeptide SEQ ID NO 2683.	148	96
1472	AAO12550	Homo sapiens	Human polypeptide SEQ ID NO 26442.	255	62
1472	AAG03600	Homo sapiens	Human secreted protein, SEQ ID NO: 7681.	245	86
1472	gi7770239	Homo sapiens	PRO2831	136	71
1473	AAW49698	Homo sapiens	Human Notch3 protein.	99	55
1473	gi3065951	Homo sapiens	Notch3	99	55
1473	AAR67242	Homo sapiens	Huaman glial cell growth factor	96	42

Table 2B

275

SEQ ID	Hit ID	Species	Description	S score	% Identity
			heregulin segment E.		
1474	gi22002433	Homo sapiens	p150 target of rapamycin (TOR)-scaffold protein containing WD-repeats	323	54
1474	gi21979456	Homo sapiens	raptor	323	54
1474	gi22002435	Mus musculus	p150 target of rapamycin (TOR)-scaffold protein containing WD-repeats	318	53
1475	AAB38280	Homo sapiens	Human secreted protein sequence encoded by gene 20 SEQ ID NO:136.	289	96
1476	gi7160973	Homo sapiens	VNN3 protein	2254	87
1476	AAM40410	Homo sapiens	Human polypeptide SEQ ID NO 3555.	2249	87
1476	gi6102996	Mus musculus	Vanin-3	1857	71
1477	gi16877231	Homo sapiens	Similar to RIKEN cDNA 2700019D07 gene	997	100
1478	gi340201	Homo sapiens	voltage-dependent anion channel	479	74
1478	gi15277577	Homo sapiens	voltage-dependent anion channel 2	479	74
1478	gi5114261	Homo sapiens	voltage-dependent anion channel isoform 2	479	74
1479	ABB08511	Homo sapiens	Human protein FH2-13.	192	48
1479	AAB94622	Homo sapiens	Human protein sequence SEQ ID NO:15476.	131	69
1479	ABJ03741	Homo sapiens	Human ovary specific protein SEQ ID NO: 183.	128	37
1480	AAB93446	Homo sapiens	Human protein sequence SEQ ID NO:12690.	204	100
1480	AAM80227	Homo sapiens	Human protein SEQ ID NO 3873.	204	100
1480	AAM79243	Homo sapiens	Human protein SEQ ID NO 1905.	204	100
1482	AAW42095	Homo sapiens	Human Rab protein A (HRABA).	192	80
1482	gi1457954	Homo sapiens	Rab22b	192	80
1482	gi10179679	Homo sapiens	small GTPase RAB22B	192	80
1483	ABG62142	Homo sapiens	Human prostate specific polypeptide #75.	170	65
1483	AAE24060	Homo sapiens	Human prostate specific protein (PSP) #3.	170	65
1483	gi7959778	Homo sapiens	PRO1546	163	65
1484	AAU69407	Homo sapiens	Lung small cell carcinoma antigen #1.	586	85
1484	ABB03602	Homo sapiens	Human musculoskeletal system related polypeptide SEQ ID NO 1549.	215	69
1484	gi6979921	Drosophila melanogaster	RhoGTPase	159	32
1485	AAU99098	Homo sapiens	Human glycosyl transferase 47169.	391	67
1485	gi18676612	Homo sapiens	FLJ00205 protein	391	67
1485	gi14150450	Rattus norvegicus	UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase T9	343	60
1486	AAO11830	Homo sapiens	Human polypeptide SEQ ID NO 25722.	283	82
1486	AAO12132	Homo sapiens	Human polypeptide SEQ ID NO 26024.	280	81
1487	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:121.	1249	85

Table 2B

276

SEQ ID	Hit ID	Species	Description	S score	% Identity
1488	AAAY27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	390	50
1488	gi13560707	Homo sapiens	group XIII secreted phospholipase A2	389	50
1488	gi12698926	Mus musculus	group XIII secreted phospholipase A2	311	80
1489	gi2995442	Homo sapiens	UDPGal:GlcNAc b1,4 galactosyltransferase	1920	90
1489	gi4520136	Homo sapiens	beta-1,4-galactosyltransferase II	1920	90
1489	gi3132896	Homo sapiens	beta-1,4-galactosyltransferase	1877	88
1490	AAO11284	Homo sapiens	Human polypeptide SEQ ID NO 25176.	165	69
1490	AAO06074	Homo sapiens	Human polypeptide SEQ ID NO 19966.	164	63
1490	AAO02002	Homo sapiens	Human polypeptide SEQ ID NO 15894.	157	64
1491	AAB93371	Homo sapiens	Human protein sequence SEQ ID NO:12521.	487	98
1491	AAM41334	Homo sapiens	Human polypeptide SEQ ID NO 6265.	487	98
1491	AAM41333	Homo sapiens	Human polypeptide SEQ ID NO 6264.	487	98
1492	AAE19435	Homo sapiens	Human 27803 (a member of human adenylate kinase family).	341	65
1492	AAU29224	Homo sapiens	Human PRO polypeptide sequence #201.	197	57
1493	gi21410587	Homo sapiens	similar to RIKEN cDNA 2310041H06	818	100
1493	gi6855513	Gallus gallus	syndesmos	551	60
1493	gi18034388	Mus musculus	syndesmos	537	58
1494	gi20987486	Homo sapiens	similar to B cell phosphoinositide 3-kinase adaptor	805	92
1494	gi12082725	Mus musculus	B cell phosphoinositide 3-kinase adaptor	331	62
1494	gi12082811	Gallus gallus	B cell phosphoinositide 3-kinase adaptor	205	48
1495	AAB64482	Homo sapiens	Human secreted protein sequence encoded by gene 13 SEQ ID NO:120.	273	100
1495	AAB65921	Homo sapiens	Human secreted protein SEQ ID NO: 61.	273	100
1495	AAB95728	Homo sapiens	Human protein sequence SEQ ID NO:18606.	273	100
1496	gi56463	Rattus norvegicus	gp210 (AA 1-1886)	7172	72
1496	gi6650678	Mus musculus	nuclear pore membrane glycoprotein POM210	7114	71
1496	gi18676550	Homo sapiens	FLJ00172 protein	5038	95
1497	AAAY40432	Homo sapiens	Amino acid sequence of vk65.15, a human VK gene fragment.	607	100
1497	AAW62185	Homo sapiens	Human DNA vkappa65.15 fragment.	607	100
1497	AAW41147	Homo sapiens	Human vkappa65.15 fragment.	607	100
1498	gi431857	Homo sapiens	delta 4-3-oxosteroid 5 beta-reductase	460	76
1498	gi11640835	Homo sapiens	5-beta steroid reductase	460	76
1498	gi5689216	Oryctolagus	delta4-3-oxosteroid 5beta-reductase	442	70

Table 2B

277

SEQ ID	Hit ID	Species	Description	S score	% Identity
		cuniculus			
1500	gi21748478	Homo sapiens	FLJ00264 protein	3399	99
1500	AAB93159	Homo sapiens	Human protein sequence SEQ ID NO:12081.	2209	64
1500	AAB58796	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 504.	1456	64
1502	AAB58202	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 540.	427	100
1502	AAU79340	Homo sapiens	Human caveolin-1 (Cav-1).	345	100
1502	AAE22092	Homo sapiens	Human caveolin-1 protein.	345	100
1503	gi3115996	Homo sapiens	dJ79C4.1.2 (Homeobox protein PMX-1 (PHOX1) isoform 2)	410	97
1503	gi3115995	Homo sapiens	dJ79C4.1.1 (Homeobox protein PMX-1 (PHOX1) isoform 1)	410	97
1503	gi460125	Mus musculus	homeobox protein	405	96
1504	gi8163762	Homo sapiens	membrane cofactor protein CD46 variant	639	83
1504	AAG75528	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6292.	635	82
1504	AAB58394	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 732.	635	82
1505	AAO21802	Homo sapiens	Lung-specific amino acid sequence SEQ ID No 113.	301	61
1505	AAO07848	Homo sapiens	Human polypeptide SEQ ID NO 21740.	161	89
1505	AAO09394	Homo sapiens	Human polypeptide SEQ ID NO 23286.	115	51
1506	AAB94891	Homo sapiens	Human protein sequence SEQ ID NO:16231.	214	65
1506	gi1335205	Homo sapiens	ORFII	209	69
1507	ABB44566	Homo sapiens	Human wound healing related polypeptide SEQ ID NO 23.	1058	100
1507	gi11414896	Homo sapiens	nucleoporin	1058	100
1507	gi1184173	Homo sapiens	nucleoporin 98	1058	100
1508	AAM79518	Homo sapiens	Human protein SEQ ID NO 3164.	1062	96
1508	AAM78534	Homo sapiens	Human protein SEQ ID NO 1196.	1062	96
1508	ABB11951	Homo sapiens	Human K/Cl cotransporter homologue, SEQ ID NO:2321.	1062	96
1509	AAB21033	Homo sapiens	Human nucleic acid-binding protein, NuABP-37.	293	81
1509	gi6691968	Homo sapiens	dJ148M19.1 (zinc finger protein)	293	81
1509	gi4096339	Homo sapiens	zinc finger protein	293	81
1511	gi8809808	Mus musculus	KRAB zinc finger protein	329	39
1511	AAM40475	Homo sapiens	Human polypeptide SEQ ID NO 5406.	322	79
1511	AAM38689	Homo sapiens	Human polypeptide SEQ ID NO 1834.	322	79
1512	gi15029737	Mus musculus	complement component 2 (within H-2S)	89	26
1512	gi3986766	Mus musculus	C2	89	26
1512	gi192435	Mus musculus	complement component C2	86	26
1513	AAB60489	Homo sapiens	Human cell cycle and proliferation protein CCYPR-37, SEQ ID NO:37.	1039	100
1513	AAB38417	Homo sapiens	Fragment of human secreted protein	1039	100

Table 2B

278

SEQ ID	Hit ID	Species	Description	S score	% Identity
			encoded by gene 5 clone HCGMF16.		
1513	gi4337460	Homo sapiens	neuroblastoma-amplified protein	1039	100
1514	gi1621611	Homo sapiens	TRAF family member-associated NF-kB activator TANK	2026	87
1514	AAAY88237	Homo sapiens	Human I-TRAF protein.	1996	86
1514	AAW27163	Homo sapiens	Human TRAF inhibitor protein I-TRAF.	1996	86
1515	ABP47963	Homo sapiens	Human polypeptide SEQ ID NO 393.	239	100
1515	AAG67216	Homo sapiens	Amino acid sequence of human Parkin-Associated Protein 1 (PAP1).	239	100
1515	gi18490728	Mus musculus	synaptotagmin-like 3	142	60
1516	AAB93432	Homo sapiens	Human protein sequence SEQ ID NO:12661.	1322	86
1516	AAW88438	Homo sapiens	Disease associated protein kinase DAPK-7.	1322	86
1516	gi5931569	Homo sapiens	H91620p	1322	86
1517	gi5931821	Homo sapiens	dJ228H13.3 (zinc finger protein)	2360	100
1517	ABP51457	Homo sapiens	Human MDDT SEQ ID NO 479.	1497	100
1517	ABP51362	Homo sapiens	Human MDDT SEQ ID NO 384.	1497	100
1518	AAU80035	Homo sapiens	Beta 2 microglobulin (beta2M)/HFE monochain.	602	100
1518	ABP43154	Homo sapiens	Human ovarian antigen HVVCG93, SEQ ID NO:4286.	602	100
1518	AAAY44412	Homo sapiens	Wild type human beta-2 microglobulin.	602	100
1520	AAO07974	Homo sapiens	Human polypeptide SEQ ID NO 21866.	158	68
1520	AAO07413	Homo sapiens	Human polypeptide SEQ ID NO 21305.	149	65
1520	AAO05994	Homo sapiens	Human polypeptide SEQ ID NO 19886.	121	56
1521	gi11386113	Homo sapiens	FKSG25	2457	94
1521	gi13548673	Homo sapiens	SCOT-t	2451	93
1521	gi20988313	Homo sapiens	3-oxoacid CoA transferase 2	2448	94
1522	AAM42400	Homo sapiens	Human polypeptide SEQ ID NO 133.	215	56
1522	AAM92851	Homo sapiens	Human digestive system antigen SEQ ID NO: 2200.	215	56
1522	gi17862572	Drosophila melanogaster	LD38503p	179	31
1523	gi3419880	Homo sapiens	MDC/ADAM11	3883	91
1523	AAR75352	Homo sapiens	Human fetal brain MDC protein.	3878	91
1523	gi5736634	Mus musculus	ADAM11	3658	87
1524	gi1109782	Homo sapiens	protein-tyrosine phosphatase	2517	98
1524	gi1781037	Mus musculus	neuronal tyrosine threonine phosphatase 1	2083	79
1524	AAM25744	Homo sapiens	Human protein sequence SEQ ID NO:1259.	987	45
1525	AAG75396	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6160.	442	100
1525	AAB53529	Homo sapiens	Human colon cancer antigen protein sequence SEQ ID NO:1069.	442	100
1525	AAR84194	Homo sapiens	Human A2b adenosine receptor.	442	100
1526	ABB44587	Homo sapiens	Human wound healing related polypeptide SEQ ID NO 44.	336	75
1526	gi184448	Homo sapiens	transcription factor	336	75

Table 2B  
279

SEQ ID	Hit ID	Species	Description	S score	% Identity
1526	gi183930	Homo sapiens	helix-loop-helix protein	336	75
1527	AAB43940	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1385.	1241	78
1527	gi2198557	Homo sapiens	spermidine aminopropyltransferase	1190	78
1527	gi14602778	Homo sapiens	spermine synthase	1190	78
1528	ABP41860	Homo sapiens	Human ovarian antigen HTPHO72, SEQ ID NO:2992.	1189	62
1528	AAW21949	Homo sapiens	E6-binding protein E6-BPSD7.	1189	62
1528	gi469885	Homo sapiens	EF-hand protein	1189	62
1529	AAU83711	Homo sapiens	Human PRO protein, Seq ID No 240.	606	100
1529	AAB84327	Homo sapiens	Amino acid sequence of a human lyase and associated protein HLYAP-2.	606	100
1529	AAU12301	Homo sapiens	Human PRO6079 polypeptide sequence.	606	100
1530	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:121.	3094	92
1530	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	3094	92
1530	ABJ03747	Homo sapiens	Human ovary specific protein SEQ ID NO: 189.	3085	91
1531	gi13477353	Homo sapiens	Similar to rabphilin 3A-like (without C2 domains)	428	51
1531	ABB06134	Homo sapiens	Human NS protein sequence SEQ ID NO:226.	427	51
1531	gi5596433	Homo sapiens	candidate tumor suppressor protein NOC2	330	87
1532	AAU81224	Homo sapiens	Human lung cancer protein, Seq ID No 57.	2576	86
1532	gi1531645	Rattus norvegicus	C2-HC type zinc finger protein r-MyT3	1913	66
1532	gi2914751	Rattus norvegicus	neural zinc finger factor 3; NZF-3	1908	66
1533	AAE23388	Homo sapiens	Human intracellular signalling (INTSIG-15) protein.	643	39
1533	ABB78745	Homo sapiens	Human Tigger 1 transposase protein sequence.	438	71
1533	AAO01214	Homo sapiens	Human polypeptide SEQ ID NO 15106.	366	65
1534	AAU27733	Homo sapiens	Human full-length polypeptide sequence #58.	3947	90
1534	AAM78731	Homo sapiens	Human protein SEQ ID NO 1393.	3943	90
1534	AAB95860	Homo sapiens	Human protein sequence SEQ ID NO:18924.	3934	89
1535	gi15928572	Mus musculus	Similar to leucine rich repeat (in FLII) interacting protein 2	103	66
1535	gi17389307	Homo sapiens	Similar to leucine rich repeat (in FLII) interacting protein 2	97	63
1535	gi5257201	Homo sapiens	LRR FLI-I interacting protein 2	97	63
1536	AAG02639	Homo sapiens	Human secreted protein, SEQ ID NO: 6720.	160	71
1536	AAG02753	Homo sapiens	Human secreted protein, SEQ ID NO: 6834.	141	60
1536	gi7959778	Homo sapiens	PRO1546	140	60
1537	ABG47195	Homo sapiens	Human peptide encoded by genome-	507	100

Table 2B

280

SEQ ID	Hit ID	Species	Description	S score	% Identity
			derived single exon probe SEQ ID 36860.		
1537	AAM38436	Homo sapiens	Peptide #12473 encoded by probe for measuring placental gene expression.	507	100
1537	AAM22055	Homo sapiens	Peptide #8489 encoded by probe for measuring cervical gene expression.	507	100
1538	gi3133291	Homo sapiens	mitogen activated protein kinase activated protein kinase	853	85
1538	gi17512453	Mus musculus	MAP kinase-activated protein kinase 5	850	84
1538	gi2911813	Mus musculus	mitogen-activated protein kinase-activated protein kinase	850	84
1539	ABG38678	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 28343.	339	100
1539	AAM04612	Homo sapiens	Peptide #3294 encoded by probe for measuring breast gene expression.	339	100
1539	AAM29402	Homo sapiens	Peptide #3439 encoded by probe for measuring placental gene expression.	339	100
1540	gi13162677	Homo sapiens	GLUT4 enhancer factor	1499	65
1540	gi14627122	Homo sapiens	dJ583P15.5.1 (Glut4 enhancer factor (isoform 3) )	1466	65
1540	AAB58934	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 642.	1395	99
1541	ABB90747	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 226.	1606	98
1541	ABB50291	Homo sapiens	Collagen type III alpha-1 ovarian tumour marker protein, SEQ ID NO:72.	1606	98
1541	ABG60248	Homo sapiens	Human ovarian antigen #10.	1606	98
1542	AAE01436	Homo sapiens	Human gene 1 encoded secreted protein HWLFJ10, SEQ ID NO:91.	1622	99
1542	AAE01515	Homo sapiens	Human gene 1 encoded secreted protein fragment, SEQ ID NO:172.	1618	98
1542	AAE01464	Homo sapiens	Human gene 1 encoded secreted protein HWLFJ10, SEQ ID NO:119.	1618	98
1543	gi186043	Homo sapiens	immunoglobulin light chain variable region	546	89
1543	gi219886	Homo sapiens	Ig kappa light chain	543	89
1543	AAV96298	Homo sapiens	Human IGFAM-10 immunoglobulin.	540	89
1544	AAE19183	Homo sapiens	Human protease, PRTS-20 protein.	3530	100
1544	AAU74761	Homo sapiens	Human protease PRTS-21 protein sequence.	2219	63
1544	AAO20514	Homo sapiens	Protein of APP related human homologue hCP201588.	2037	98
1545	gi6693836	Rattus norvegicus	SNIP-b	4413	78
1545	gi6693834	Rattus norvegicus	SNIP-a	4410	81
1545	gi3098418	Mus musculus	P140	4130	79
1546	gi179433	Homo sapiens	biglycan	2043	95
1546	AAB85043	Homo sapiens	Human biglycan protein sequence.	1894	89
1546	gi6960459	Homo sapiens	biglycan	1894	89
1547	gi20987689	Homo sapiens	Similar to allantoicase	1332	64



Table 2B  
281

SEQ ID	Hit ID	Species	Description	S score	% Identity
1547	gi9255889	Mus musculus	allantoicase	1095	53
1547	gi14718648	Homo sapiens	allantoicase	1071	66
1548	AAM79289	Homo sapiens	Human protein SEQ ID NO 1951.	203	78
1548	AAG03927	Homo sapiens	Human secreted protein, SEQ ID NO: 8008.	203	78
1548	AAB57110	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1688.	203	78
1549	gi1838956	Bos taurus	capping protein, beta3 isoform	156	100
1549	gi2239063	Bos taurus	actin-binding protein CP3	156	100
1549	gi595257	Homo sapiens	F-actin capping protein beta subunit	156	100
1550	gi3370998	Homo sapiens	BAI1-associated protein 1	526	91
1550	gi15278193	Homo sapiens	MAGI-1C beta	526	91
1550	gi15278186	Homo sapiens	MAGI-1A	526	91
1551	ABG61604	Homo sapiens	Human DPRP-2 splice variant #3.	2850	95
1551	ABG61602	Homo sapiens	Human DPRP-2 splice variant #1.	2850	95
1551	ABG61592	Homo sapiens	Human DPPIV related serine protease DPRP-2.	2850	95
1552	AAG73867	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:4631.	322	36
1552	AAB58391	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 729.	322	36
1552	gi296667	Homo sapiens	ornithine decarboxylase (ODC)	322	36
1553	AAW96153	Homo sapiens	Human FADD-interacting protein (FIP).	168	100
1553	gi16118555	Homo sapiens	ELMO1	168	100
1553	gi21594742	Mus musculus	Similar to engulfment and cell motility 1, ced-12 homolog (C. elegans)	168	100
1554	ABB89950	Homo sapiens	Human polypeptide SEQ ID NO 2326.	244	81
1554	AAU86135	Homo sapiens	Human PRO274 polypeptide.	244	81
1554	AAB93352	Homo sapiens	Human protein sequence SEQ ID NO:12476.	244	81
1556	AAM38652	Homo sapiens	Human polypeptide SEQ ID NO 1797.	1838	85
1556	gi4530437	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	1645	78
1556	AAM40438	Homo sapiens	Human polypeptide SEQ ID NO 5369.	1611	77
1557	AAB64943	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:121.	2327	89
1557	AAB38012	Homo sapiens	Human secreted protein encoded by gene 3 clone HNHCT15.	2322	89
1558	gi11066463	Rattus norvegicus	RhoGEF glutamate transport modulator GTRAP48	3794	72
1558	gi19387126	Mus musculus	guanine nucleotide exchange factor	1051	37
1558	AAB90743	Homo sapiens	Human CW420_2 protein sequence SEQ ID 186.	1049	50
1559	gi18033747	Homo sapiens	myosin IIIB	1032	66
1559	AAE24138	Homo sapiens	Human kinase (PKIN)-9 protein.	964	65
1559	AAU03552	Homo sapiens	Human protein kinase #52.	964	65
1560	AAM79228	Homo sapiens	Human protein SEQ ID NO 1890.	3488	64
1560	AAM39300	Homo sapiens	Human polypeptide SEQ ID NO 2445.	3467	64

Table 2B

282

SEQ ID	Hit ID	Species	Description	S score	% Identity
1560	AAM41086	Homo sapiens	Human polypeptide SEQ ID NO 6017.	3458	63
1561	AAB59019	Homo sapiens	Breast and ovarian cancer associated antigen protein sequence SEQ ID 727.	158	47
1561	ABG35360	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 25025.	131	60
1561	AAM01340	Homo sapiens	Peptide #22 encoded by probe for measuring human breast gene expression.	131	60
1562	gi1504002	Homo sapiens	similar to a human major CRK-binding protein DOCK180.	2549	82
1562	gi13195147	Mus musculus	HCH	2457	79
1562	AAW03515	Homo sapiens	Human DOCK180 protein.	1636	53
1563	AAB52017	Homo sapiens	Human secreted protein sequence encoded by gene 6 SEQ ID NO:66.	266	100
1563	AAM91576	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:19169.	77	37
1564	AAU10690	Homo sapiens	Human agouti polypeptide.	362	97
1564	gi540073	Homo sapiens	agouti precursor	362	97
1564	gi8953446	Homo sapiens	dJ785G19.3 (agouti (mouse)-signaling protein)	362	97
1565	AAM79853	Homo sapiens	Human protein SEQ ID NO 3499.	2302	82
1565	AAM78869	Homo sapiens	Human protein SEQ ID NO 1531.	2302	82
1565	gi619877	Homo sapiens	hydroxymethylglutaryl-CoA synthase	2302	82
1566	AAM41373	Homo sapiens	Human polypeptide SEQ ID NO 6304.	356	100
1566	AAM39587	Homo sapiens	Human polypeptide SEQ ID NO 2732.	356	100
1566	AAU12276	Homo sapiens	Human PRO6001 polypeptide sequence.	154	41
1567	AAU27723	Homo sapiens	Human full-length polypeptide sequence #48.	232	80
1567	AAO10211	Homo sapiens	Human polypeptide SEQ ID NO 24103.	219	84
1567	AAO01262	Homo sapiens	Human polypeptide SEQ ID NO 15154.	218	74
1568	gi11595428	Homo sapiens	dJ702J19.1 (glycine cleavage system protein H (aminomethyl carrier))	862	91
1568	gi219671	Homo sapiens	hydrogen carrier protein precursor	856	91
1568	gi12653985	Homo sapiens	glycine cleavage system protein H (aminomethyl carrier)	856	91
1569	AAO14211	Homo sapiens	Human transporter and ion channel TRICH-28.	228	59
1569	AAG77968	Homo sapiens	Human ion channel protein IC23949.	217	57
1569	AAM34851	Homo sapiens	Peptide #8888 encoded by probe for measuring placental gene expression.	155	100
1570	gi532505	Homo sapiens	bile acid CoA: Amino acid N-acyltransferase	2049	99
1570	gi16306985	Homo sapiens	bile acid Coenzyme A: amino acid N-acyltransferase (glycine N-choloyltransferase)	2049	99
1570	gi15215152	Mus musculus	Similar to bile acid-Coenzyme A	1410	68

Table 2B

283

SEQ ID	Hit ID	Species	Description	S score	% Identity
			dehydrogenase: amino acid n-acyltransferase		
1571	gi15487674	Homo sapiens	OSBP-related protein 1; ORP1	204	100
1571	gi17529999	Homo sapiens	oxysterol-binding protein-like protein OSBPL1B	204	100
1571	AAG89290	Homo sapiens	Human secreted protein, SEQ ID NO: 410.	198	97
1572	AAO17114	Homo sapiens	Human Gli3 protein SEQ ID NO: 21.	1105	62
1572	gi6102812	Homo sapiens	GLI3 protein	1105	62
1572	gi183248	Homo sapiens	DNA-binding protein	1105	62
1573	gi18447471	Drosophila melanogaster	RE41571p	177	47
1573	gi9965400	murine herpesvirus 72	membrane virion glycoprotein 150	128	30
1573	gi1019435	Trypanosoma cruzi	mucin-like protein	127	36
1574	ABG38560	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 28225.	239	100
1574	AAM04517	Homo sapiens	Peptide #3199 encoded by probe for measuring breast gene expression.	239	100
1574	AAM29283	Homo sapiens	Peptide #3320 encoded by probe for measuring placental gene expression.	239	100
1575	AAU83597	Homo sapiens	Human PRO protein, Seq ID No 12.	1294	100
1575	AAB74709	Homo sapiens	Human membrane associated protein MEMAP-15.	1294	100
1575	AAU96185	Homo sapiens	Human secreted protein, SEQ ID No 87.	1121	80
1576	AAM79494	Homo sapiens	Human protein SEQ ID NO 3140.	2432	86
1576	gi1773381	Homo sapiens	APXL	2432	86
1576	AAM78510	Homo sapiens	Human protein SEQ ID NO 1172.	2428	86
1577	AAU07829	Homo sapiens	Human ARTS-1 polypeptide.	3395	76
1577	gi21315078	Homo sapiens	type 1 tumor necrosis factor receptor shedding aminopeptidase regulator	3395	76
1577	gi6381989	Homo sapiens	adipocyte-derived leucine aminopeptidase	3388	76
1578	ABG35337	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 25002.	150	75
1578	AAM01319	Homo sapiens	Peptide #1 encoded by probe for measuring human breast gene expression.	150	75
1578	AAM25964	Homo sapiens	Peptide #1 encoded by probe for measuring placental gene expression.	150	75
1579	gi7657864	Homo sapiens	match to nuclear protein, NP220; note: sequence difference at residue 58	4034	98
1579	AAV07032	Homo sapiens	Breast cancer associated antigen precursor sequence.	4023	98
1579	gi1374698	Homo sapiens	nuclear protein, NP220	4023	98
1580	ABB09439	Homo sapiens	Serp domain protein Zserp15 associated amino acid sequence.	1172	98
1580	AAW51933	Homo sapiens	Human protease nexin I type alpha.	1172	98
1580	AAW51934	Homo sapiens	Human protease nexin I type beta.	1172	98
1581	gi499184	Felis catus	neuronal protein	305	93

Table 2B  
284

SEQ ID	Hit ID	Species	Description	S score	% Identity
1581	AAB95041	Homo sapiens	Human protein sequence SEQ ID NO:16804.	284	73
1581	gi13960126	Homo sapiens	Similar to leucine-rich neuronal protein	230	75
1582	AAB80253	Homo sapiens	Human PRO293 protein.	3181	99
1582	AAAY13385	Homo sapiens	Amino acid sequence of protein PRO293.	3181	99
1582	AAU83686	Homo sapiens	Human PRO protein, Seq ID No 190.	1669	55
1583	AAB73858	Homo sapiens	Human NPM/ALK fusion protein.	268	75
1583	AAB28184	Homo sapiens	Human Nucleophosmin.	268	75
1584	AAB93492	Homo sapiens	Human protein sequence SEQ ID NO:12796.	697	100
1584	gi7302187	Drosophila melanogaster	CG7849-PA	274	40
1584	gi21626800	Drosophila melanogaster	CG7849-PB	274	40
1585	gi8572061	Bombyx mori	fibroin heavy chain Fib-H	156	34
1585	gi765323	Bombyx mori	silk fibroin heavy chain	155	39
1585	gi155999	Bombyx mori	silk fibroin	147	39
1586	gi12002682	Homo sapiens	FERM-containing protein	2019	83
1586	ABP37983	Homo sapiens	Human GS93382 protein.	531	99
1586	AAB12318	Homo sapiens	Human secreted protein encoded by gene 18 clone HE2FL70.	189	100
1587	ABP41887	Homo sapiens	Human ovarian antigen HKZBB48, SEQ ID NO:3019.	1005	90
1587	AAB95585	Homo sapiens	Human protein sequence SEQ ID NO:18250.	1005	90
1587	AAAY86208	Homo sapiens	Nuclear transport protein clone hfb030 protein sequence.	1005	90
1590	AAG02851	Homo sapiens	Human secreted protein, SEQ ID NO: 6932.	226	88
1590	AAB38280	Homo sapiens	Human secreted protein sequence encoded by gene 20 SEQ ID NO:136.	226	88
1591	AAB95593	Homo sapiens	Human protein sequence SEQ ID NO:18273.	2718	82
1591	AAU27637	Homo sapiens	Human protein AFP39158.	609	75
1591	ABG45908	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 35573.	311	100
1592	ABB05746	Homo sapiens	Human GASC1 protein SEQ ID NO:1.	780	33
1592	gi10567164	Homo sapiens	gene amplified in squamous cell carcinoma-1	780	33
1592	AAM39339	Homo sapiens	Human polypeptide SEQ ID NO 2484.	777	33
1593	AAAY69069	Homo sapiens	Amino acid sequence of a human reduced tropoelastin derivative.	2887	80
1593	gi182020	Homo sapiens	elastin	2887	80
1593	AAO17360	Homo sapiens	Human elastin.	2870	79
1595	gi291854	Homo sapiens	aminopeptidase A	1289	91
1595	gi1518865	Sus scrofa	aminopeptidase A	1127	80
1595	gi7673021	Rattus norvegicus	aminopeptidase A	1103	77
1596	AAU71952	Homo sapiens	Human bone marrow tissue	356	58

Table 2B  
285

SEQ ID	Hit ID	Species	Description	S score	% Identity
			polypeptide #30.		
1597	gi556769	Homo sapiens	inositol 1,4,5-triphosphate 5-phosphatase	342	92
1597	ABB77441	Homo sapiens	Human tumour marker protein Li9-1.	339	91
1597	gi21262190	Homo sapiens	CTCL tumor antigen Li9-1	339	91
1598	ABB08405	Homo sapiens	Alpha1,6-fucosyl transferase amino acid sequence.	367	100
1598	AAW22125	Homo sapiens	Human alpha 1-6 fucosyltransferase.	367	100
1598	gi7638407	Bos taurus	6-alpha-L-fucosyltransferase	367	100
1599	AAM90773	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:18366.	905	97
1599	AAU74354	Homo sapiens	Human cytoskeleton-associated protein (CYSKP) #25.	740	43
1599	AAB93267	Homo sapiens	Human protein sequence SEQ ID NO:12300.	149	46
1600	AAB48140	Homo sapiens	Human TANGO 209 variant 2 polypeptide.	372	80
1600	AAB48139	Homo sapiens	Human TANGO 209 variant 1 polypeptide.	372	80
1600	AAB48107	Homo sapiens	Human TANGO 209 polypeptide.	372	80
1601	gi6941623	Homo sapiens	dJ469D22.1 (Rhesus blood group, CcEe antigens)	268	76
1601	AAB03339	Homo sapiens	Human RhCe protein.	252	100
1601	AAB03338	Homo sapiens	Human RhCE protein.	252	100
1602	AAB94196	Homo sapiens	Human protein sequence SEQ ID NO:14530.	784	79
1602	AAB93449	Homo sapiens	Human protein sequence SEQ ID NO:12696.	784	79
1602	AAB97178	Homo sapiens	Human EF-chiral calcium-binding protein 96.	784	79
1603	gi13096814	Mus musculus	RIKEN cDNA 4932442K08 gene	165	72
1603	gi11493928	Homo sapiens	homeodomain-interacting protein kinase 2	74	20
1604	gi1321596	Homo sapiens	immunoglobulin lambda-chain subgroup II	378	63
1604	AAG76014	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6778.	362	97
1604	gi21619848	Homo sapiens	Similar to immunoglobulin lambda joining 3	344	92
1605	AAB95234	Homo sapiens	Human protein sequence SEQ ID NO:17375.	2336	92
1605	AAB27239	Homo sapiens	Human EXMAD-17 SEQ ID NO: 17.	2334	99
1605	gi19263740	Mus musculus	axotrophin	1935	79
1606	AAB92702	Homo sapiens	Human protein sequence SEQ ID NO:11102.	3250	92
1606	gi20451682	Homo sapiens	JEMMA protein	3250	92
1606	AAB93188	Homo sapiens	Human protein sequence SEQ ID NO:12140.	3088	91
1607	AAB31693	Homo sapiens	Amino acid sequence of viral encoded semaphorin protein receptor.	1014	79
1607	AAB70131	Homo sapiens	Human VESPR.	1014	79
1607	AAAY13462	Homo sapiens	Viral-encoded semaphorin protein receptor (VESPR) polypeptide.	1014	79
1608	AAW29683	Homo sapiens	Human Na-K-2Cl cotransporter NKCC2.	1295	99

Table 2B

286

SEQ ID	Hit ID	Species	Description	S score	% Identity
1608	gi1373425	Homo sapiens	bumetanide-sensitive Na-K-2Cl cotransporter	1295	99
1608	gi516001	Oryctolagus cuniculus	bumetanide-sensitive Na-K-Cl cotransport protein splice isoform F	1244	95
1610	AAB88388	Homo sapiens	Human membrane or secretory protein clone PSEC0131.	923	100
1610	gi20810493	Homo sapiens	Similar to RIKEN cDNA 2810417M05 gene	920	99
1610	AAE21272	Homo sapiens	Human gene 16 encoded secreted protein fragment, SEQ ID NO:138.	432	100
1611	AAY53753	Homo sapiens	Amino acid sequence of the MMSC2 protein.	269	98
1611	AAY04732	Homo sapiens	Protein containing PDZ domain from clone 38-2-1a.	269	98
1611	AAY04734	Homo sapiens	Protein containing PDZ domain from clone 38-2-1c.	269	98
1612	AAR85912	Homo sapiens	Oncostatin M receptor-beta subunit.	1138	92
1612	gi1794211	Homo sapiens	oncostatin-M specific receptor beta subunit	1138	92
1612	gi15012082	Homo sapiens	Similar to oncostatin M receptor	1138	92
1613	ABP42361	Homo sapiens	Human ovarian antigen HNOKG34, SEQ ID NO:3493.	406	82
1613	AAB48966	Homo sapiens	Human heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1).	406	82
1613	AAW55828	Homo sapiens	Human heterogeneous nuclear ribonucleoprotein core protein A1.	406	82
1614	ABB07649	Homo sapiens	Human LOR-1 protein.	898	99
1614	AAB00077	Homo sapiens	Human lysyl oxidase related protein (Lor).	898	99
1614	gi1890108	Homo sapiens	lysyl oxidase-related protein	898	99
1615	gi12803157	Homo sapiens	COX15 (yeast) homolog, cytochrome c oxidase assembly protein	163	48
1615	gi3603230	Homo sapiens	cytochrome oxidase assembly factor	163	48
1615	gi7619702	Homo sapiens	bA483F11.2.1 (COX15 (yeast) homolog, cytochrome c oxidase assembly protein (isoform 1))	163	48
1616	gi7672979	Homo sapiens	glucosidase II beta subunit	2553	84
1616	ABB50266	Homo sapiens	Protein kinase C substrate 80K-H ovarian tumour marker protein, #20.	2536	84
1616	gi1293640	Homo sapiens	protein kinase C substrate 80K-H	2536	84
1617	AAY07076	Homo sapiens	Renal cancer associated antigen precursor sequence.	355	100
1617	gi1022888	Oryctolagus cuniculus	protein phosphatase 2A0 B' subunit beta4 isoform	355	100
1617	gi1022886	Oryctolagus cuniculus	protein phosphatase 2A0 B' regulatory subunit beta3 isoform	355	100
1618	gi14010930	Haemophilus influenzae Rd; similar to AAC22226.1 (PID:g1573555) [Homo sapiens]	similar to transcription accessory protein (tex)	1868	100
1618	AAB92758	Homo sapiens	Human protein sequence SEQ ID NO:11220.	1863	99
1618	ABB89105	Homo sapiens	Human polypeptide SEQ ID NO	1231	95

Table 2B

287

SEQ ID	Hit ID	Species	Description	S score	% Identity
			1481.		
1619	AAW64262	Homo sapiens	Human neutrophil elastase.	628	100
1619	AAP80335	Homo sapiens	Sequence of serine protease (SP) of human myeloid cellorigin and leader peptide.	628	100
1619	gi296665	Homo sapiens	serine protease	628	100
1620	gi9965989	Homo sapiens	calcineurin A catalytic subunit gamma isoform	1648	97
1620	gi258001	Homo sapiens	calcineurin A catalytic subunit; calmodulin-dependent protein phosphatase catalytic subunit; CaM-PrP catalytic subunit	1637	96
1620	gi200466	Mus musculus	phosphoprotein phosphatase	1471	85
1621	gi1778538	Escherichia coli	anaerobic carrier for c4, dicarboxylates	739	100
1621	gi1786839	Escherichia coli K12	transport of dicarboxylates	739	100
1621	gi12513520	Escherichia coli O157:H7 EDL933	transport of dicarboxylates	739	100
1622	AAE15635	Homo sapiens	Human G-protein coupled receptor-5 (GCREC-5) protein.	271	94
1622	AAB66272	Homo sapiens	Human TANGO 378 SEQ ID NO: 29.	271	94
1622	AAB82487	Homo sapiens	Human secretin-like receptor Zgpr1 splice variant.	271	94
1623	gi1787320	Escherichia coli K12	homolog of Salmonella P-ring of flagella basal body	749	87
1623	gi4062658	Escherichia coli	Flagellar basal body P-ring protein precursor	749	87
1623	gi12514621	Escherichia coli O157:H7 EDL933	homolog of Salmonella P-ring of flagella basal body	748	86
1624	gi13360031	Escherichia coli O157:H7	allantoinase	492	69
1624	gi2735238	Escherichia coli	GlxB3	491	70
1624	gi1773192	Escherichia coli	similar to S. cerevisiae dal1	491	70
1625	gi606368	Escherichia coli	aspartate semialdehyde dehydrogenase	1478	93
1625	gi3859587	Shigella sonnei	aspartate semialdehyde dehydrogenase	1478	93
1625	gi1789841	Escherichia coli K12	aspartate-semialdehyde dehydrogenase	1478	93
1626	gi1788006	Escherichia coli K12	phenylalanine tRNA synthetase, beta-subunit	735	94
1626	gi12515724	Escherichia coli O157:H7 EDL933	phenylalanine tRNA synthetase, beta-subunit	735	94
1626	gi1742806	Escherichia coli	Phenylalanine--tRNA ligase (EC 6.1.1.20) b chain	735	94
1627	AAB49502	Homo sapiens	Clone HYASC03.	310	98
1627	gi22137373	Mus musculus	similar to RIKEN cDNA 2810051A14 gene	240	95
1627	gi20071228	Mus musculus	RIKEN cDNA 2810051A14 gene	151	51
1628	gi17431382	Ralstonia	PROBABLE TRANSCRIPTION	357	37

Table 2B  
288

SEQ ID	Hit ID	Species	Description	S score	% Identity
		solanacearum	REGULATOR PROTEIN		
1628	gi21114950	Xanthomonas campestris pv. campestris str. ATCC 33913	transcriptional regulator	331	38
1629	gi311422	Escherichia coli	ORF-2	282	100
1630	gi732874	Neisseria meningitidis	IgA1 protease	101	33
1630	gi600255	Gallus gallus	caldesmon	99	23
1631	gi1788756	Escherichia coli K12	PEP-protein phosphotransferase system enzyme I	785	96
1631	gi12516792	Escherichia coli O157:H7 EDL933	PEP-protein phosphotransferase system enzyme I	785	96
1631	gi1799835	Escherichia coli	PHOSPHOENOLPYRUVATE-PROTEIN PHOSPHOTRANSFERASE (EC 2.7.3.9) (PHOSPHOTRANSFERASE SYSTEM, ENZYME I).	785	96
1632	gi887820	Escherichia coli	UUG start; possible frameshift at end?	810	72
1632	gi466651	Escherichia coli	No definition line found	343	82
1632	gi21957797	Yersinia pestis KIM	acridine efflux pump	250	56
1633	AAB92950	Homo sapiens	Human protein sequence SEQ ID NO:11629.	290	100
1633	gi21627153	Drosophila melanogaster	CG8155-PA	82	39
1633	gi3850257	Coltivirus JKT-7043	Vp9	78	40
1634	gi12580933	Homo sapiens	dJ505P2.1.1 (ribonuclease 6 precursor)	442	62
1634	AAV21852	Homo sapiens	Human signal peptide-containing protein (SIGP) (clone ID 2652271).	437	61
1634	AAW75103	Homo sapiens	Human secreted protein encoded by gene 47 clone HMCBP63.	437	61
1635	gi18146756	Homo sapiens	limkain beta 2	1620	100
1635	AAV86509	Homo sapiens	Human gene 70-encoded protein fragment, SEQ ID NO:424.	696	100
1635	AAV86510	Homo sapiens	Human gene 70-encoded protein fragment, SEQ ID NO:425.	436	100
1636	gi8096340	Homo sapiens	RERE	513	71
1636	gi4689163	Caenorhabditis elegans	EGL-27	83	23
1636	gi21913127	Caenorhabditis elegans	C. elegans EGL-27 protein (corresponding sequence C04A2.3b)	83	23
1637	AAU69417	Homo sapiens	Lung small cell carcinoma antigen #11.	514	94
1637	AAB85481	Homo sapiens	Human 23553 sulfatase polypeptide.	514	94
1637	AAM25714	Homo sapiens	Human protein sequence SEQ ID NO:1229.	514	94
1638	AAM43540	Homo sapiens	Human polypeptide SEQ ID NO 218.	741	100
1638	AAM38682	Homo sapiens	Human polypeptide SEQ ID NO 1827.	741	100



Table 2B  
289

SEQ ID	Hit ID	Species	Description	S score	% Identity
1638	AAY42750	Homo sapiens	Human calcium binding protein 1 (CaBP-1).	741	100
1639	AAB94175	Homo sapiens	Human protein sequence SEQ ID NO:14484.	908	86
1639	AAM40361	Homo sapiens	Human polypeptide SEQ ID NO 3506.	908	86
1639	gi15215085	Mus musculus	Similar to COP9 (constitutive photomorphogenic), subunit 7b (Arabidopsis)	900	85
1640	gi1529	Oryctolagus cuniculus	calcium channel BI-2	95	30
1640	gi2653364	Bovine herpesvirus type 1.1	immediate-early transactivator protein (cell nucleus)	91	25
1640	gi291536	Bovine herpesvirus 1	BICP4	91	25
1641	AAB26105	Homo sapiens	Human DAN/Cerberus-related protein 6 (hDCR6) #1.	2012	100
1641	AAE17089	Homo sapiens	Human osteolevin protein.	758	100
1641	ABB07209	Homo sapiens	Human cloaked-2 polypeptide sequence.	758	100
1642	AAU29026	Homo sapiens	Human PRO polypeptide sequence #3.	279	94
1642	AAY99458	Homo sapiens	Human PRO193 amino acid sequence SEQ ID NO:410.	279	94
1642	AAY36125	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 510.	279	94
1643	ABG60166	Homo sapiens	Human DITHP polypeptide #224.	426	76
1643	AAG76121	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6885.	426	76
1643	AAM38668	Homo sapiens	Human polypeptide SEQ ID NO 1813.	426	76
1644	AAU84339	Homo sapiens	Protein HPV16E1Bind differentially expressed in breast cancer tissue.	514	82
1644	gi2232019	Homo sapiens	HPV16 E1 protein binding protein	514	82
1644	gi12653271	Homo sapiens	thyroid hormone receptor interactor 13	514	82
1645	AAU74752	Homo sapiens	Human protease PRTS-12 protein sequence.	3063	99
1645	ABG34074	Homo sapiens	Human Pro peptide #45.	2954	98
1645	AAB50936	Homo sapiens	ADAM protein #2.	2814	100
1646	AAO20516	Homo sapiens	Protein of APP related human homologue hCP50592.	273	98
1646	gi1232077	Homo sapiens	huMCM2	273	98
1646	gi13544066	Homo sapiens	Similar to mini chromosome maintenance deficient 2 (S. cerevisiae)	273	98
1647	AAG75416	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6180.	655	74
1647	AAB57016	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1594.	655	74
1647	gi14328059	Homo sapiens	argininosuccinate synthetase	655	74
1648	gi7960207	Oncorhynchus mykiss	vitelline envelope protein alpha	252	43
1648	gi15384295	Mycoplasma	variable surface lipoprotein Vsp422-	177	34

Table 2B  
290

SEQ ID	Hit ID	Species	Description	S score	% Identity
		bovis	8		
1648	gi2281750	Salmo salar	eggshell protein	164	38
1649	AAM79049	Homo sapiens	Human protein SEQ ID NO 1711.	345	78
1650	gi7018384	Homo sapiens	dJ193N13.1 (mannosidase, alpha, class 1A, member 1)	1026	98
1650	gi15929672	Mus musculus	mannosidase 1, alpha	909	88
1650	gi474280	Mus musculus	mannosyl-oligosaccharide alpha-1,2-mannosidase	909	88
1651	AAU10551	Homo sapiens	Human A259 polypeptide.	835	71
1651	ABP47883	Homo sapiens	Human polypeptide SEQ ID NO 313.	835	71
1651	AAB50085	Homo sapiens	Human A259.	835	71
1652	gi4512295	Homo sapiens	immunoglobulin heavy chain variable region	619	100
1652	AAR66320	Homo sapiens	Human immunoglobulin variable heavy chain #26.	613	99
1652	gi296657	Homo sapiens	Ig heavy chain gene variable region V(12G-1)	613	99
1653	AAE24079	Homo sapiens	Human MDPK protein.	185	85
1653	AAE24150	Homo sapiens	Human kinase (PKIN)-21 protein.	181	97
1653	AAU03501	Homo sapiens	Human protein kinase #1.	181	97
1654	AAU75784	Homo sapiens	Human protein phosphatase 2 (PP2) protein sequence.	667	92
1654	AAE04841	Homo sapiens	Human SGP039 phosphatase polypeptide.	667	92
1654	gi8954030	117	Contains similarity to protein phosphatase 2C from Arabidopsis thaliana gb AF085279. It contains a protein phosphatase 2C domain PF00481	42	
1655	AAU96179	Homo sapiens	Human secreted protein, SEQ ID No 81.	86	27
1655	gi1794219	Homo sapiens	150 kDa oxygen-regulated protein ORP150	82	25
1655	AAB01381	Homo sapiens	Neuron-associated protein.	76	41
1656	ABB84869	Homo sapiens	Human PRO1079 protein sequence SEQ ID NO:106.	404	78
1656	ABB95475	Homo sapiens	Human angiogenesis related protein PRO1079 SEQ ID NO: 106.	404	78
1656	AAB75372	Homo sapiens	Human secreted protein #31.	404	78
1657	AAM41200	Homo sapiens	Human polypeptide SEQ ID NO 6131.	332	97
1657	AAM39414	Homo sapiens	Human polypeptide SEQ ID NO 2559.	332	97
1657	gi31867	Homo sapiens	N-acetylglucosamine-6-sulphatase	332	97
1658	AAW25154	Homo sapiens	Human disulphide epimerase like enzyme, EP52.	2281	100
1658	gi1136743	Homo sapiens	human P5	2281	100
1658	gi12654931	Homo sapiens	protein disulfide isomerase-related protein	2281	100
1659	AAM41859	Homo sapiens	Human polypeptide SEQ ID NO 6790.	357	71
1659	gi6694278	Homo sapiens	cell recognition molecule Caspr2	357	71
1659	gi13624214	Homo sapiens	contactin-associated protein 2	357	71
1660	AAO11734	Homo sapiens	Human polypeptide SEQ ID NO 25626.	141	100

Table 2B  
291

SEQ ID	Hit ID	Species	Description	S score	% Identity
1661	AAG74841	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:5605.	568	87
1661	gi4588087	Homo sapiens	PTH-responsive osteosarcoma B1 protein	568	87
1661	gi20987880	Mus musculus	Similar to PTH-responsive osteosarcoma B1 protein	536	86
1662	gi4754907	Homo sapiens	histone deacetylase 4	2891	77
1662	gi14495171	Gallus gallus	histone deacetylase-4	2525	68
1662	AAB49957	Homo sapiens	Human histone deacetylase HDAC-4.	2508	79
1663	ABB94075	Homo sapiens	Human secreted protein SEQ ID NO: 118.	183	100
1663	ABB94056	Homo sapiens	Human secreted protein SEQ ID NO: 99.	183	100
1663	ABB94030	Homo sapiens	Human secreted protein SEQ ID NO: 73.	183	100
1664	ABB37187	Homo sapiens	Peptide #4693 encoded by human foetal liver single exon probe.	196	100
1664	gi13310486	Homo sapiens	C2H2 zinc finger protein	91	31
1664	gi15159543	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_L_2143p	88	25
1665	AAB32388	Homo sapiens	Human secreted protein sequence encoded by gene 18 SEQ ID NO:74.	359	100
1665	AAAY91419	Homo sapiens	Human secreted protein sequence encoded by gene 6 SEQ ID NO:140.	82	36
1665	gi21708132	Homo sapiens	Similar to LOC149473	82	31
1666	gi219615	Homo sapiens	preproelafin	621	100
1666	gi299841	Homo sapiens	pre-elafin	621	100
1666	gi190338	Homo sapiens	elafin precursor	621	100
1667	gi18072031	Homo sapiens	zinc finger protein 328	397	78
1667	AAU27687	Homo sapiens	Human full-length polypeptide sequence #12.	270	66
1667	AAM79885	Homo sapiens	Human protein SEQ ID NO 3531.	259	58
1669	AAG03136	Homo sapiens	Human secreted protein, SEQ ID NO: 7217.	179	75
1669	AAU85309	Homo sapiens	G-coupled olfactory receptor #170.	176	68
1669	AAU24689	Homo sapiens	Human olfactory receptor AOLFR188.	176	68
1670	AAAY21852	Homo sapiens	Human signal peptide-containing protein (SIGP) (clone ID 2652271).	375	83
1670	AAAY48563	Homo sapiens	Human breast tumour-associated protein 24.	375	83
1670	AAW75103	Homo sapiens	Human secreted protein encoded by gene 47 clone HMCBP63.	375	83
1671	gi6572310	Homo sapiens	cB13C9.1 (protein 239AB)	395	98
1671	AAU69508	Homo sapiens	Human purified secretory polypeptide #77.	348	60
1671	gi7374112	Homo sapiens	239AB	347	80
1672	AAG63163	Homo sapiens	Amino acid sequence of carcinoembryonic antigen-like polypeptide.	2187	99
1672	AAR54714	Homo sapiens	Carcinoembryonic antigen CEA-(c).	456	30
1672	AAB43688	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1133.	455	31
1673	AAU74354	Homo sapiens	Human cytoskeleton-associated	1163	34

Table 2B

292

SEQ ID	Hit ID	Species	Description	S score	% Identity
			protein (CYSKP) #25.		
1673	AAB93267	Homo sapiens	Human protein sequence SEQ ID NO:12300.	463	33
1675	AAM06070	Homo sapiens	Peptide #4752 encoded by probe for measuring breast gene expression.	259	100
1676	ABP43066	Homo sapiens	Human ovarian antigen HVCAE01, SEQ ID NO:4198.	292	58
1676	ABP42762	Homo sapiens	Human ovarian antigen HOVJU75, SEQ ID NO:3894.	292	58
1676	AAG01285	Homo sapiens	Human secreted protein, SEQ ID NO: 5366.	292	58
1677	gi458226	Homo sapiens	a gene isolated from a CpG island between STS and KAL	1249	97
1677	AAG00737	Homo sapiens	Human secreted protein, SEQ ID NO: 4818.	531	94
1677	gi9663151	Homo sapiens	transport-secretion protein 2.1 (TTS-2.1)	421	35
1678	gi1806102	Homo sapiens	T cell receptor beta chain	444	100
1678	AAB68370	Homo sapiens	Human beta-chain variable region of T cell receptors (Vbeta17).	440	98
1678	AAW76993	Homo sapiens	Human T cell receptor beta chain variable region protein V-beta 17.	440	98
1679	AAB92624	Homo sapiens	Human protein sequence SEQ ID NO:10919.	455	100
1679	ABG39897	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 29562.	155	69
1679	AAM05720	Homo sapiens	Peptide #4402 encoded by probe for measuring breast gene expression.	155	69
1680	gi10186503	Homo sapiens	sialic acid-specific acetyltransferase II	907	68
1680	gi10242345	Homo sapiens	sialic acid-specific 9-O-acetyltransferase I	907	68
1680	gi5917657	Mus musculus	cytosolic sialic acid 9-O-acetyltransferase	732	55
1681	gi46540	Staphylococcus aureus	coagulase precursor	82	25
1681	gi20338653	Staphylococcus aureus	coagulase	82	25
1681	gi21328339	Caenorhabditis elegans	C. elegans RBF-1 protein (corresponding sequence F37A4.7d)	80	31
1682	AAO07844	Homo sapiens	Human polypeptide SEQ ID NO 21736.	100	70
1682	AAO07413	Homo sapiens	Human polypeptide SEQ ID NO 21305.	91	46
1682	gi17861952	Drosophila melanogaster	LD01947p	87	53
1683	AAW99574	Homo sapiens	Human early placental insulin-like (EPIL) polypeptide.	349	100
1683	AAW26926	Homo sapiens	Human insulin-like 4 protein.	349	100
1683	AAW69168	Homo sapiens	Zins1 protein.	349	100
1684	gi23365	Homo sapiens	17-HSD protein (AA 1 - 328)	1082	70
1684	gi181951	Homo sapiens	estradiol 17 beta-dehydrogenase	1082	70
1684	gi177127	Homo sapiens	17-beta-hydroxysteroid dehydrogenase	1082	70
1685	ABB84963	Homo sapiens	Human PRO4356 protein sequence	1336	99

Table 2B  
293

SEQ ID	Hit ID	Species	Description	S score	% Identity
			SEQ ID NO:294.		
1685	ABB95569	Homo sapiens	Human angiogenesis related protein PRO4356 SEQ ID NO: 294.	1336	99
1685	AAB31206	Homo sapiens	Amino acid sequence of human polypeptide PRO4356.	1336	99
1686	ABB80870	Homo sapiens	Type II TGFbeta receptor (RII) sequence.	2130	88
1686	AAB82996	Homo sapiens	Human transforming growth factor-beta receptor RII.	2130	88
1686	AAY59453	Homo sapiens	Human Transforming growth factor-beta II protein sequence.	2130	88
1687	ABB79162	Homo sapiens	Human VPAC1 receptor protein SEQ ID NO:2.	386	76
1687	ABB79161	Homo sapiens	Human VPAC1 receptor protein SEQ ID NO:1.	386	76
1687	AAB71878	Homo sapiens	Human VIPR seven transmembrane domain.	386	76
1688	AAB84663	Homo sapiens	Amino acid sequence of human tyrosine kinase protein Yes.	520	90
1688	AAY24421	Homo sapiens	Human yes1 protein.	520	90
1688	gi181268	Homo sapiens	cellular yes-1 protein	520	90
1690	AAO14938	Homo sapiens	Human UFD2-associated protein 1.	1386	76
1690	AAM00761	Homo sapiens	Human bone marrow protein, SEQ ID NO: 124.	1386	76
1690	gi13516467	Homo sapiens	homzygously deleted in neuroblastoma-1/UFD2	1386	76
1691	AAB47977	Homo sapiens	BCY5.	334	98
1691	gi2463632	Homo sapiens	monocarboxylate transporter homologue MCT6	326	96
1691	gi9246437	Staphylococcus aureus	fntA-like protein	81	24
1692	AAB95143	Homo sapiens	Human protein sequence SEQ ID NO:17163.	394	98
1692	AAB93175	Homo sapiens	Human protein sequence SEQ ID NO:12114.	394	98
1692	AAB58175	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 513.	394	98
1693	AAG65906	Homo sapiens	Amino acid sequence of GSK gene Id 37967.	3025	92
1693	AAY67598	Homo sapiens	Human adipose tissue protein #1.	501	40
1693	gi3786312	Homo sapiens	extracellular matrix protein	501	40
1694	ABB83488	Homo sapiens	Human cytoskeleton-associated protein, CSAP-17.	216	93
1694	gi3002588	Mus musculus	Plenty of SH3s; POSH	216	93
1694	gi18676610	Homo sapiens	FLJ00204 protein	206	86
1695	gi12018147	Chlamydomonas reinhardtii	vegetative cell wall protein gp1	132	30
1695	gi17945382	Drosophila melanogaster	RE17165p	130	28
1695	gi1209103	Rattus norvegicus	atrophin-1 related protein	130	25
1696	AAB43791	Homo sapiens	Human cancer associated protein sequence SEQ ID NO:1236.	353	98
1696	AAM79791	Homo sapiens	Human protein SEQ ID NO 3437.	152	77
1696	AAM78807	Homo sapiens	Human protein SEQ ID NO 1469.	152	77

Table 2B  
294

SEQ ID	Hit ID	Species	Description	S score	% Identity
1697	AAE22224	Homo sapiens	Human 5685C6 protein.	542	98
1697	ABG40551	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 30216.	516	98
1697	AAM18652	Homo sapiens	Peptide #5086 encoded by probe for measuring cervical gene expression.	516	98
1698	AAM39274	Homo sapiens	Human polypeptide SEQ ID NO 2419.	1596	79
1698	AAB29653	Homo sapiens	Human membrane-associated protein HUMAP-10.	1596	79
1698	gi9858855	Homo sapiens	HPT protein	1596	79
1699	gi6841138	Homo sapiens	HSPC099	275	100
1699	AAB93037	Homo sapiens	Human protein sequence SEQ ID NO:11816.	130	23
1699	AAB61308	Homo sapiens	Human transcriptional regulator protein #8.	125	22
1700	AAM93959	Homo sapiens	Human polypeptide, SEQ ID NO: 4164.	1182	99
1700	AAB36587	Homo sapiens	Human FLEXHT-9 protein sequence SEQ ID NO:9.	590	53
1700	AAB93652	Homo sapiens	Human protein sequence SEQ ID NO:13161.	586	54
1701	gi30264	Homo sapiens	cystatin D	392	100
1701	gi16116526	Homo sapiens	yX60D10.1 (cystatin D)	387	100
1701	AAO15149	Homo sapiens	Human cystatin D protein sequence.	380	98
1702	gi1374754	Staphylothermus marinus	tetrabrachion	72	27
1703	AAE17127	Homo sapiens	Human GnT-V protein.	628	77
1703	AAB48911	Homo sapiens	Human beta-1,6-N-acetylglucosamine transferase (GnT-V).	628	77
1703	AAB83010	Homo sapiens	Human GnT-V.	628	77
1704	AAAY68736	Homo sapiens	Short chain alcohol dehydrogenase-related molecule ScRM-2 cDNA.	1249	81
1704	AAAY69295	Homo sapiens	A human human protein, designated HSPC021 (CBFAIG06).	1249	81
1704	AAB58463	Homo sapiens	Lung cancer associated polypeptide sequence SEQ ID 801.	1249	81
1705	gi15294065	Ictalurus punctatus	40S ribosomal protein S26-2	212	58
1705	ABP42978	Homo sapiens	Human ovarian antigen HPDWD69, SEQ ID NO:4110.	208	58
1705	AAG76127	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6891.	208	58
1706	gi6580428	Homo sapiens	IkappaBR	2759	95
1706	gi14250636	Homo sapiens	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 2	2606	86
1706	gi746415	Homo sapiens	I kappa BR	1655	71
1707	AAM79219	Homo sapiens	Human protein SEQ ID NO 1881.	325	64
1708	gi5901529	Homo sapiens	C2H2 type Kruppel-like zinc finger protein splice variant b	565	99
1708	gi5901527	Homo sapiens	C2H2 type Kruppel-like zinc finger protein	317	33
1708	gi15029916	Mus musculus	Similar to Zinc finger protein 118	296	31

Table 2B  
295

SEQ ID	Hit ID	Species	Description	S score	% Identity
1709	AAU85415	Homo sapiens	Human protein NOV13.	953	98
1709	AAU72895	Homo sapiens	Human metalloprotease partial protein sequence #7.	925	100
1709	AAU74750	Homo sapiens	Human protease PRTS-10 protein sequence.	925	100
1710	gi12862392	Mus musculus	D86	2863	64
1710	ABG37531	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 27196.	1687	100
1710	AAM03553	Homo sapiens	Peptide #2235 encoded by probe for measuring breast gene expression.	1687	100
1711	gi1787337	acyl-carrier-protein synthase II [Escherichia coli K12]	3-oxoacyl-	315	91
1711	gi12514639	acyl-carrier-protein synthase II [Escherichia coli O157:H7 EDL933]	3-oxoacyl-	315	91
1711	gi664870	Escherichia coli	beta-ketoacyl-acyl carrier protein synthase II	315	91
1712	AAG75407	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6171.	1258	82
1712	AAR89952	Homo sapiens	Insulin-like growth factor binding protein-3.	1232	80
1712	AAU85512	Homo sapiens	Clone #19095 (L549S) of lung tumour protein.	1228	80
1713	AAB94696	Homo sapiens	Human protein sequence SEQ ID NO:15673.	710	97
1713	AAB99892	Homo sapiens	Human RNA helicase gene helicain C protein sequence SEQ ID NO:6.	710	97
1713	AAB99891	Homo sapiens	Human RNA helicase gene helicain B protein sequence SEQ ID NO:4.	710	97
1714	gi15384740	Homo sapiens	paralemm-2	1652	100
1714	gi15384742	Homo sapiens	Palm2-AKAP2 fusion protein	1577	100
1714	gi14041780	Homo sapiens	AKAP-2 protein	410	73
1715	gi13021825	Homo sapiens	polymerase	575	48
1715	gi3600067	Homo sapiens	polymerase	572	48
1715	gi1780973	Human endogenous retrovirus K	pol protein	572	48
1717	gi35825	Homo sapiens	pregnancy zone protein	6592	88
1717	gi579594	Homo sapiens	alpha 2-macroglobulin 690-740	4985	66
1717	AAR11749	Homo sapiens	Human alpha-2 macroglobulin bait region mutant.	4976	66
1718	gi13366277	Homo sapiens	dJ998H6.1 (ortholog of rat PB-Cadherin)	1167	93
1718	gi4760578	Mus musculus	PB-Cadherin	1028	84
1718	gi1398912	Rattus norvegicus	short type PB-cadherin	1022	84
1720	AAE06588	Homo sapiens	Human protein having hydrophobic domain, HP10778.	687	100
1720	AAM40979	Homo sapiens	Human polypeptide SEQ ID NO 5910.	687	100

Table 2B

296

SEQ ID	Hit ID	Species	Description	S score	% Identity
1720	gi15072402	Raja erinacea	organic solute transporter alpha	357	45
1721	AAG81345	Homo sapiens	Human AFP protein sequence SEQ ID NO:208.	839	62
1721	gi16359082	Homo sapiens	Similar to RIKEN cDNA 2810049G06 gene	839	62
1721	AAB93797	Homo sapiens	Human protein sequence SEQ ID NO:13560.	836	62
1722	gi871883	Homo sapiens	lanosterol 14-demethylase	2180	99
1722	gi1809225	Homo sapiens	lanosterol 14-demethylase (cytochrome p450)	2180	99
1722	gi1698396	Homo sapiens	lanosterol 14-demethylase cytochrome P450	2180	99
1723	ABG41541	Homo sapiens	Human peptide encoded by genome-derived single exon probe SEQ ID 31206.	232	100
1723	AAM32019	Homo sapiens	Peptide #6056 encoded by probe for measuring placental gene expression.	232	100
1723	AAM71727	Homo sapiens	Human bone marrow expressed probe encoded protein SEQ ID NO: 32033.	232	100
1724	AAV53040	Homo sapiens	Human secreted protein clone kj320_1 protein sequence SEQ ID NO:86.	2480	100
1724	gi3510639	Rattus norvegicus	UDP-GalNAc:polypeptide N-acetylglactosaminyltransferase T5	1345	59
1724	gi6688167	Homo sapiens	GalNAc-T5	1082	100
1725	ABP41917	Homo sapiens	Human ovarian antigen H6EDF71, SEQ ID NO:3049.	1329	97
1725	AAG75406	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:6170.	1329	97
1725	gi396176	Homo sapiens	antigenic surface determinant OA3	1329	97
1726	AAU79036	Homo sapiens	Human SHPS-1 (not defined) receptor.	858	98
1726	AAW40481	Homo sapiens	Human SH2 binding protein.	858	98
1726	AAW49909	Homo sapiens	Signal regulatory protein 4 (SIRP4).	858	98
1727	gi2707601	Homo sapiens	synaptophysin	1656	100
1727	gi15928723	Mus musculus	synaptophysin	1585	94
1727	gi57326	Rattus norvegicus	synaptophysin (AA 1-307)	1577	95
1728	gi15590682	Homo sapiens	histone deacetylase 9a	628	96
1728	gi15590680	Homo sapiens	histone deacetylase 9	628	96
1728	gi12060992	Mus musculus	MEF2-interacting transcription repressor MITR	625	95
1729	gi21430596	Drosophila melanogaster	RE16431p	891	39
1729	gi14164377	Mus musculus	Type II membrane protein of ER~mouse gene similar to alpha-mannosidase	350	33
1729	gi1504008	Homo sapiens	Containing ATP/GTP-binding site motif A(P-loop): Similar to C.elegans protein(P1:CEC47E128);Similar to Mouse alpha-mannosidase(P1:B54407)	346	33
1730	gi7861753	Mus musculus	GABA-A receptor epsilon-like	412	32



Table 2B  
297

SEQ ID	Hit ID	Species	Description	S score	% Identity
			subunit		
1730	gi11610622	Rattus norvegicus	GABA-A epsilon subunit splice variant	411	32
1730	gi7861743	Rattus norvegicus	GABA-A receptor epsilon-like subunit	411	32
1731	AAW80135	Homo sapiens	Human recombinant neurokinin-2 (NK-2) receptor protein.	1743	95
1731	gi189222	Homo sapiens	neurokinin-2 receptor	1743	95
1731	gi189135	Homo sapiens	neurokinin A receptor	1742	95
1732	AAB75594	Homo sapiens	Human secreted protein sequence encoded by gene 37 SEQ ID NO:148.	678	99
1732	AAB80437	Homo sapiens	Gene #20 associated peptide #1.	381	98
1732	AAM78175	Homo sapiens	Human bone marrow expressed probe encoded protein SEQ ID NO: 38481.	365	100
1733	ABB84853	Homo sapiens	Human PRO1120 protein sequence SEQ ID NO:74.	293	100
1733	ABB90378	Homo sapiens	Human polypeptide SEQ ID NO 2754.	293	100
1733	ABB95459	Homo sapiens	Human angiogenesis related protein PRO1120 SEQ ID NO: 74.	293	100
1734	ABB07527	Homo sapiens	Human drug metabolizing enzyme (DME) (ID: 5643401CD1).	1652	79
1734	ABB07515	Homo sapiens	Human drug metabolizing enzyme (DME) (ID: 8097779CD1).	867	73
1734	gi13161409	Mus musculus	family 4 cytochrome P450	718	58
1735	AAM40183	Homo sapiens	Human polypeptide SEQ ID NO 3328.	1322	67
1735	AAY71159	Homo sapiens	Human phosphodiesterase interacting protein, myomegalin.	1322	67
1735	gi4761644	Rattus norvegicus	myomegalin	886	44
1736	AAM94312	Homo sapiens	Human reproductive system related antigen SEQ ID NO: 2970.	500	71
1736	ABJ03726	Homo sapiens	Human ovary specific protein SEQ ID NO: 168.	298	41
1736	gi8439396	HERV-H/env62	envelope protein	292	40
1737	AAY76177	Homo sapiens	Human secreted protein encoded by gene 54.	288	100
1738	AAY92075	Homo sapiens	Human DKR-4.	759	100
1738	AAB08875	Homo sapiens	Amino acid sequence of a human Dickkopf (Dkk)-4 protein.	759	100
1738	AAW73017	Homo sapiens	Human cysteine-rich secreted protein CRSP-2.	759	100
1739	ABB97828	Homo sapiens	Human secretory polypeptide (SPTM) 80.	2007	86
1739	ABB90159	Homo sapiens	Human polypeptide SEQ ID NO 2535.	1547	98
1739	gi18642980	Homo sapiens	GTPase	1524	61
1740	AAB49278	Homo sapiens	Protein encoded by zsig81 cDNA fragment.	755	94
1740	AAU29276	Homo sapiens	Human PRO polypeptide sequence #253.	755	94
1740	gi12003127	Eremothecium	GTPase activating protein BEM2	83	28

Table 2B  
298

SEQ ID	Hit ID	Species	Description	S score	% Identity
		gossypii			
1741	gi21707232	Homo sapiens	similar to choline transporter-like protein	1637	76
1741	AAB65196	Homo sapiens	Human PRO1115 (UNQ558) protein sequence SEQ ID NO:177.	1636	76
1741	AAB87541	Homo sapiens	Human PRO1115.	1636	76
1742	AAG66142	Homo sapiens	Human PAS Kinase (PASK) polypeptide.	2227	99
1742	AAM79231	Homo sapiens	Human protein SEQ ID NO 1893.	2227	99
1742	AAM79230	Homo sapiens	Human protein SEQ ID NO 1892.	2227	99
1743	gi13879899	Mycobacterium tuberculosis CDC1551	PPE family protein	102	27
1743	gi2653311	Bovine herpesvirus type 1.1	very large virion protein (tegument)	99	27
1743	gi1491621	Bovine herpesvirus 1	UL36	99	27
1744	gi5931718	Chlamydomonas reinhardtii	1-alpha dynein heavy chain	2126	56
1744	gi9409781	Chlamydomonas reinhardtii	1 beta dynein heavy chain	1121	34
1744	gi514215	Chlamydomonas reinhardtii	dynein beta heavy chain	1070	33
1745	AAB53088	Homo sapiens	Human angiogenesis-associated protein PRO328, SEQ ID NO:132.	1972	89
1745	AAB80260	Homo sapiens	Human PRO328 protein.	1972	89
1745	AAU12351	Homo sapiens	Human PRO328 polypeptide sequence.	1972	89
1746	AAU29172	Homo sapiens	Human PRO polypeptide sequence #149.	730	68
1746	AAV99398	Homo sapiens	Human PRO1301 (UNQ667) amino acid sequence SEQ ID NO:212.	730	68
1746	AAM38651	Homo sapiens	Human polypeptide SEQ ID NO 1796.	728	68
1747	AAE21056	Homo sapiens	Human drug metabolising enzyme (DME-14) protein.	111	59
1748	AAE03560	Homo sapiens	Human differentially expressed kidney cDNA 22360 encoded protein.	1140	88
1748	AAM42434	Homo sapiens	Human kidney related polypeptide SEQ ID NO 303.	466	98
1748	AAM99619	Homo sapiens	Human excretory related polypeptide SEQ ID NO 356.	466	98
1749	AAE22911	Homo sapiens	Human transporter and ion channel (TRICH) 10.	2309	100
1749	gi13506805	Homo sapiens	thymic stromal co-transporter	2309	100
1749	gi13506808	Mus musculus	thymic stromal co-transporter	1782	77
1750	gi18157547	Mus musculus	pecanex-like 3	7005	89
1750	gi15076843	Homo sapiens	pecanex-like protein 1	4359	51
1750	gi13171105	Takifugu rubripes	pecanex	4118	58
1751	AAV06603	Homo sapiens	Retinoblastoma interacting protein GluT1*.	411	100
1751	gi825504	Homo sapiens	glutamate transporter	411	100

Table 2B  
299

SEQ ID	Hit ID	Species	Description	S score	% Identity
1751	gi825663	Homo sapiens	GLAST1	411	100
1752	gi2078518	Homo sapiens	neogenin	593	100
1752	gi641966	Gallus gallus	neogenin	591	98
1752	gi1785999	Rattus norvegicus	neogenin	586	97
1753	ABB53278	Homo sapiens	Human polypeptide #18.	315	26
1753	ABB53277	Homo sapiens	Human polypeptide #17.	315	26
1753	AAP94014	Homo sapiens	Carcinoembryonic cell surface antigen.	225	25
1754	AAO21681	Homo sapiens	Human secreted protein SEQ ID No 23.	443	95
1754	AAB75375	Homo sapiens	Human secreted protein #34.	443	95
1754	AAB88603	Homo sapiens	Human hydrophobic domain containing protein clone HP10770 #127.	443	95
1755	AAE22906	Homo sapiens	Human transporter and ion channel (TRICH) 5.	1973	90
1755	AAE16350	Homo sapiens	Human tetracycline transporter like-like protein, POLY14.	1969	90
1755	gi2506078	Mus musculus	tetracycline transporter-like protein	1961	90
1756	AAB88469	Homo sapiens	Human membrane or secretory protein clone PSEC0027.	1065	98
1756	AAE06608	Homo sapiens	Human protein having hydrophobic domain, HP10798.	1065	98
1756	AAM40347	Homo sapiens	Human polypeptide SEQ ID NO 3492.	1065	98
1757	gi8925284	Homo sapiens	phosphatidylinositol polyphosphate 5-phosphatase type IV	2599	91
1757	gi9295353	Mus musculus	inositol polyphosphate 5-phosphatase	1989	78
1757	gi5360761	Rattus norvegicus	pharbin	1928	77
1758	gi395207	Bos taurus	potassium channel (BGK5)	753	82
1758	gi186669	Homo sapiens	potassium channel	750	82
1758	gi304652	Canis familiaris	delayed rectifier K+ channel	746	81
1759	AAB65058	Homo sapiens	Gene #7 associated peptide #12.	908	93
1759	AAB64999	Homo sapiens	Human secreted protein #7.	908	93
1759	AAU01099	Homo sapiens	Gene 35 Human secreted protein homologous amino acid sequence.	814	91
1760	AAU11384	Homo sapiens	Human T2R61 (hT2R61) polypeptide.	802	100
1760	gi20336531	Homo sapiens	candidate taste receptor T2RP1	794	99
1760	ABB06836	Homo sapiens	Human nGPCR-Seq1048 protein sequence SEQ ID NO:105.	718	79
1761	AAAY36115	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 500.	514	87
1761	AAW64556	Homo sapiens	Human osteocarcinoma cell line U-2 OS clone HP10305 protein.	514	87
1761	gi14250122	Homo sapiens	uncharacterized hematopoietic stem/progenitor cells protein MDS029	514	87
1762	AAG78575	Homo sapiens	Human SLC5A3 amino acid sequence.	3621	100
1762	AAB47976	Homo sapiens	BCW2.	3621	100
1762	gi2739094	Homo sapiens	sodium/myo-inositol cotransporter	3621	100
1763	ABB80578	Homo sapiens	Human sbg618069LRR protein #2.	3076	99

Table 2B  
300

SEQ ID	Hit ID	Species	Description	S score	% Identity
1763	ABB80577	Homo sapiens	Human sbg618069LRR protein #1.	1572	95
1763	AAW84596	Homo sapiens	Amino acid sequence of the human Tango-79 protein.	1209	44
1764	ABB81460	Homo sapiens	Human aggrecanase MDTS8 protein SEQ ID NO:2.	3392	89
1764	gi19171150	Homo sapiens	ADAMTS18 protein	3364	89
1764	AAU72893	Homo sapiens	Human metalloprotease partial protein sequence #5.	2762	90
1765	gi1695682	Homo sapiens	hepatic triglyceride lipase	152	64
1765	gi32498	Homo sapiens	precursor (AA -23 to 476)	152	64
1765	gi339593	Homo sapiens	triglyceride lipase	152	64
1766	AAM79459	Homo sapiens	Human protein SEQ ID NO 3105.	1493	100
1766	AAM78475	Homo sapiens	Human protein SEQ ID NO 1137.	1493	100
1766	ABB11930	Homo sapiens	Human secreted protein homologue, SEQ ID NO:2300.	1493	100
1767	AAM47914	Homo sapiens	Human lysophosphatidic aminoacyl transferase 42.	324	90
1767	AAE15296	Homo sapiens	Human LPAAT delta protein.	324	90
1767	AAB65188	Homo sapiens	Human PRO1016 (UNQ499) protein sequence SEQ ID NO:156.	324	90
1768	AAE23757	Homo sapiens	Human metabotropic glutamate (mGluR4) receptor protein.	926	99
1768	AAR82658	Homo sapiens	Human mGluR4.	926	99
1768	gi1160183	Homo sapiens	metabotropic glutamate receptor type 4	926	99
1769	AAM41363	Homo sapiens	Human polypeptide SEQ ID NO 6294.	2370	98
1769	AAE17500	Homo sapiens	Human secretion and trafficking protein-9 (SAT-9).	2207	99
1769	AAM39577	Homo sapiens	Human polypeptide SEQ ID NO 2722.	2207	99
1770	AAB73512	Homo sapiens	Human transferase HTFS-19, SEQ ID NO:19.	264	45
1770	AAB56487	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1065.	264	45
1770	gi14249942	Homo sapiens	Similar to RIKEN cDNA 061008P16 gene	264	45
1771	gi7678873	Homo sapiens	vascular cadherin-2	5369	99
1771	gi7407150	Homo sapiens	protocadherin 12	5369	99
1771	gi8164037	Homo sapiens	vascular endothelial cadherin 2	5369	99
1772	AAM93947	Homo sapiens	Human polypeptide, SEQ ID NO: 4138.	231	93
1772	AAM93886	Homo sapiens	Human polypeptide, SEQ ID NO: 4013.	231	93
1772	AAM93857	Homo sapiens	Human polypeptide, SEQ ID NO: 3949.	231	93
1773	gi1335205	Homo sapiens	ORFII	270	48
1773	AAB94891	Homo sapiens	Human protein sequence SEQ ID NO:16231.	263	62
1773	AAG73650	Homo sapiens	Human colon cancer antigen protein SEQ ID NO:4414.	263	58
1774	AAB93885	Homo sapiens	Human protein sequence SEQ ID NO:13815.	1088	85
1774	AAM93980	Homo sapiens	Human stomach cancer expressed polypeptide SEQ ID NO 29.	1088	85

Table 2B  
301

SEQ ID	Hit ID	Species	Description	S score	% Identity
1774	gi3603459	Homo sapiens	tetraspan NET-5	1088	85
1775	ABB06607	Homo sapiens	G protein-coupled receptor GPCR5 protein SEQ ID NO:24.	982	77
1775	AAG71597	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1278.	982	77
1775	gi21928553	Homo sapiens	seven transmembrane helix receptor	982	77
1776	AAB64888	Homo sapiens	Human secreted protein sequence encoded by gene 7 SEQ ID NO:66.	252	85
1776	AAB38011	Homo sapiens	Human secreted protein encoded by gene 3 clone HPJCX13.	252	85
1776	AAB94917	Homo sapiens	Human protein sequence SEQ ID NO:16357.	191	62
1777	gi182851	Homo sapiens	G0S2 protein	497	100
1777	gi1213013	Mus musculus	G0S2-like protein	377	77
1777	AAM79519	Homo sapiens	Human protein SEQ ID NO 3165.	81	26
1778	ABB89432	Homo sapiens	Human polypeptide SEQ ID NO 1808.	473	100
1778	ABP61790	Homo sapiens	Human polypeptide SEQ ID NO 144.	473	100
1778	AAAY80991	Homo sapiens	Human VAMP-2 homologue, CBCBMH06.	473	100
1779	gi5264503	Mus musculus	sif and Tiam1-like exchange factor	82	30
1779	gi9295309	frog adenovirus 1	pIIa protein	78	40
1779	gi16415263	Listeria innocua	similar to two-component sensor histidine kinase	78	27
1780	gi21693020	Homo sapiens	MHC class I antigen	1311	70
1780	gi1399321	Macaca mulatta	MHC class I antigen Mamu B*08	1307	71
1780	gi8117799	Pan troglodytes	MHC class I antigen	1302	70
1781	ABB75677	Homo sapiens	Breast protein-eukaryotic conserved gene 1 (BSTP-ECG1) protein.	754	40
1781	AAU29191	Homo sapiens	Human PRO polypeptide sequence #168.	754	40
1781	AAAY99421	Homo sapiens	Human PRO1433 (UNQ738) amino acid sequence SEQ ID NO:292.	754	40
1782	gi2343157	Homo sapiens	peroxisomal membrane protein 69	2796	89
1782	gi2706518	Homo sapiens	peroxisomal ABC-transporter	2796	89
1782	gi15215442	Homo sapiens	Similar to ATP-binding cassette, sub-family D (ALD), member 4	2788	88
1783	AAB35235	Homo sapiens	Human neurotransmitter transporter protein GC42.	3573	98
1783	AAB35236	Homo sapiens	Human glycine transporter type 1c.	3559	98
1783	gi546769	Homo sapiens	glycine transporter type 1b; GlyT-1b	3559	98
1784	AAU00017	Homo sapiens	Human Plexin-D1.	7512	90
1784	ABB11709	Homo sapiens	Human plexin-B1/SEP receptor homologue, SEQ ID NO:2079.	7467	90
1784	gi5918167	Homo sapiens	plexin-B1/SEP receptor	2120	33
1785	ABB44591	Homo sapiens	Human wound healing related polypeptide SEQ ID NO 48.	953	99
1785	gi619730	Homo sapiens	nuclear factor I	953	99
1785	gi9653290	Mus sp.	NFI-X3	953	99
1786	gi13810568	Homo sapiens	Toll-like receptor 5	4482	100
1786	ABB11795	Homo sapiens	Human Toll/IL-1R-like protein homologue, SEQ ID NO:2165.	4478	99
1786	gi3132526	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	4464	99

Table 2B  
302

SEQ ID	Hit ID	Species	Description	S score	% Identity
1787	AAB88597	Homo sapiens	Human hydrophobic domain containing protein clone HP03670 #121.	887	82
1787	AAB56473	Homo sapiens	Human prostate cancer antigen protein sequence SEQ ID NO:1051.	887	82
1787	AAB60119	Homo sapiens	Human transport protein TPPT-39.	564	75
1788	AAB80300	Homo sapiens	Human prostate cancer antigen #28.	741	89
1788	AAB80276	Homo sapiens	Human prostate cancer antigen #4.	741	89
1788	gi4929765	Homo sapiens	CGI-148 protein	741	89
1789	gi22038189	Escherichia coli	multidrug transporter	2265	100
1789	gi1736785	Escherichia coli	Acriflavin resistance protein F (EnvD protein).	2265	100
1789	gi15980819	Yersinia pestis	AcrB/AcrD/AcrF family membrane protein	1854	79
1790	gi17741602	ribose [Agrobacterium tumefaciens str. C58 (U. Washington)]	ABC transporter, membrane spanning protein	549	59
1790	gi15160166	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_L_3181p	549	59
1790	gi20515187	Thermoanaerobacter tengcongensis	Ribose/xylose/arabinose/galactoside ABC-type transport systems, permease components	340	38
1791	gi1788573	Escherichia coli K12	sn-glycerol-3-phosphate permease	1073	100
1791	gi1799587	Escherichia coli	glycerol-3-phosphate transport protein	1073	100
1791	gi41587	Escherichia coli	glycerol-3-phosphatase transporter (AA 1 - 452, glpT)	1073	100
1792	gi1790233	Escherichia coli K12	arylsulfatase	679	98
1792	gi12518665	Escherichia coli O157:H7 EDL933	arylsulfatase	679	98
1792	gi13364207	Escherichia coli O157:H7	arylsulfatase	679	98
1793	gi21959134	Yersinia pestis KIM	sulfate transporter	378	80
1793	gi9657461	Vibrio cholerae	sulfate permease family protein	267	51
1793	gi18145142	Clostridium perfringens str. 13	probable sulfate permease	264	53
1794	gi1799719	PIR Accession Number S08346 [Escherichia coli]	similar to	922	100
1794	gi15156677	Agrobacterium tumefaciens str. C58 (Cereon)	AGR_C_2926p	452	50
1795	gi11177166	Mus musculus	adhesion molecule ninjurin	228	81
1795	gi3077901	Mus musculus	ninjurin	228	81
1795	gi1644366	Rattus norvegicus	ninjurin l	228	81

Table 2B  
303

SEQ ID	Hit ID	Species	Description	S score	% Identity
1796	AAU74824	Homo sapiens	Human REPTR 7 protein.	6196	92
1796	ABB90740	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 212.	6196	92
1796	ABB90725	Homo sapiens	Human Tumour Endothelial Marker polypeptide SEQ ID NO 188.	6196	92
1797	ABP61451	Homo sapiens	Human NF-kB activating protein SEQ ID NO 55.	1331	100
1797	AAAY94343	Homo sapiens	Human cell surface receptor protein #10.	1331	100
1797	gi13938575	Homo sapiens	Similar to RIKEN cDNA 2610511E22 gene	1331	100
1798	gi606234	Escherichia coli	secY	953	100
1798	gi42989	Escherichia coli	SecY (PriA) polypeptide (aa 1-443)	953	100
1798	gi16421976	Salmonella typhimurium LT2	preprotein translocase of IISF family	950	99
1799	gi18255305	Mus musculus	p53 apoptosis effector related to Pmp22	440	53
1799	gi7582391	Mus musculus	p53 apoptosis-associated target	440	53
1799	AAM50572	Homo sapiens	Human tumour suppressor protein THW.	391	100
1800	AAU11433	Homo sapiens	Human short-chain dehydrogenase, SCDR.	531	58
1800	AAU18368	Homo sapiens	Human endocrine polypeptide SEQ ID No 323.	531	58
1800	AAU18369	Homo sapiens	Human endocrine polypeptide SEQ ID No 324.	530	59
1801	ABB12425	Homo sapiens	Human bone marrow expressed protein SEQ ID NO: 264.	1781	90
1801	AAM50318	Homo sapiens	Human membrane transporter (MTP) 33556.	1499	90
1801	AAM83805	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:11398.	1499	90
1802	gi20810074	Homo sapiens	Similar to pepsinogen 5, group I (pepsinogen A)	670	86
1802	AAB61351	Homo sapiens	Pepsin protein.	661	85
1802	AAB66589	Homo sapiens	Human pepsin.	661	85
1803	ABB04707	Homo sapiens	Human SP82 protein SEQ ID NO:2.	1113	99
1803	ABB84912	Homo sapiens	Human PRO1356 protein sequence SEQ ID NO:192.	1113	99
1803	AAU76534	Homo sapiens	Tumour-associated antigenic target protein, TAT134.	1113	99
1804	gi2641217	Oryctolagus cuniculus	anion exchanger 3 brain isoform	645	66
1804	gi476222	Homo sapiens	anion exchanger 3 brain isoform	645	66
1804	gi886256	Homo sapiens	anion exchange protein	642	66
1805	AAM25789	Homo sapiens	Human protein sequence SEQ ID NO:1304.	2659	56
1805	AAB29632	Homo sapiens	Human pollinosis-associated gene 581-encoded protein, SEQ ID NO:12.	2659	56
1805	gi18698435	Homo sapiens	pVHL-interacting deubiquitinating enzyme 1 type II	2659	56
1806	AAU29153	Homo sapiens	Human PRO polypeptide sequence #130.	1771	98

Table 2B  
304

SEQ ID	Hit ID	Species	Description	S score	% Identity
1806	AAAY99363	Homo sapiens	Human PRO1380 (UNQ717) amino acid sequence SEQ ID NO:79.	1771	98
1806	AAG67487	Homo sapiens	Amino acid sequence of a human transporter protein.	1765	98
1807	ABB76315	Homo sapiens	Human protein phosphatase PP-8 Incyte ID No. 4022502CD1.	667	73
1807	gi15778670	Mus musculus	sphingosine-1-phosphate phosphatase	151	30
1807	gi9623190	Mus musculus	sphingosine-1-phosphate phosphohydrolase	151	30
1808	AAM93947	Homo sapiens	Human polypeptide, SEQ ID NO: 4138.	231	93
1808	AAM93886	Homo sapiens	Human polypeptide, SEQ ID NO: 4013.	231	93
1808	AAM93857	Homo sapiens	Human polypeptide, SEQ ID NO: 3949.	231	93
1809	gi14575679	Homo sapiens	hemicentin	597	95
1809	gi3328186	Caenorhabditis elegans	hemicentin precursor	354	57
1809	AAU75886	Homo sapiens	Human adhesion molecule protein AD4/AAD21820.1.	185	33
1810	gi3581982	Homo sapiens	extraneuronal monoamine transporter	775	99
1810	gi14270513	Homo sapiens	organic cation transporter 3	775	99
1810	gi13699874	Mus musculus	organic cation transporter 3	720	89
1811	AAB74762	Homo sapiens	Human secreted protein sequence encoded by gene 18 SEQ ID NO:71.	266	88
1811	AAB74760	Homo sapiens	Human secreted protein sequence encoded by gene 18 SEQ ID NO:69.	266	91
1811	AAB74759	Homo sapiens	Human secreted protein sequence encoded by gene 18 SEQ ID NO:68.	266	91
1812	gi15082375	Homo sapiens	Similar to transmembrane 7 superfamily member 1 (upregulated in kidney)	912	82
1812	gi13096836	Mus musculus	Similar to transmembrane 7 superfamily member 1 (upregulated in kidney)	885	80
1812	ABB97817	Homo sapiens	Human secretory polypeptide (SPTM) 69.	432	68
1813	gi1504024	Homo sapiens	similar to Mouse finger protein(clone mkr3)(S03677):	901	37
1813	gi14549186	Mus musculus	zinc finger protein 219	636	31
1813	AAM39029	Homo sapiens	Human polypeptide SEQ ID NO 2174.	619	32
1814	AAAY53644	Homo sapiens	Protein encoded by the human longevity assurance gene 1 (LAG1).	1382	93
1814	AAR20230	Homo sapiens	hUOG-1.	1382	93
1814	gi4324468	Homo sapiens	LAG1 protein	1382	93
1815	AAU75907	Homo sapiens	Human epidermis-specific serine protease #2.	329	41
1815	gi6009515	Xenopus laevis	epidermis specific serine protease	329	41
1815	gi19353198	Mus musculus	RIKEN cDNA 2010001P08 gene	323	41
1816	AAM42401	Homo sapiens	Human polypeptide SEQ ID NO 134.	706	100
1816	gi1786232	Escherichia coli K12	K <sup>+</sup> efflux antiporter, glutathione-regulated	706	100
1816	gi21321928	Escherichia coli	Glutathione-regulated potassium-	706	100



Table 2B  
305

SEQ ID	Hit ID	Species	Description	S score	% Identity
			efflux system protein KefC (K(+)/H(+)- antiporter).		
1817	gi1657563	Escherichia coli	dioxygenase	1510	100
1817	gi1786565	Escherichia coli K12	taurine dioxygenase, 2-oxoglutarate-dependent	1510	100
1817	gi1054578	Escherichia coli	dioxygenase	1510	100
1818	gi1787550	Escherichia coli K12	homolog of Salmonella peptide transport permease protein	928	98
1818	gi12515484	Escherichia coli O157:H7 EDL933	homolog of Salmonella peptide transport permease protein	928	98
1818	gi13361335	Escherichia coli O157:H7	homolog of Salmonella peptide transport permease protein	928	98
1819	gi9652147	Homo sapiens	transmembrane-type protein tyrosine phosphatase H	5771	98
1819	gi475004	Homo sapiens	protein tyrosine phosphatase precursor	5235	93
1819	gi1321659	Rattus norvegicus	brain-enriched membrane-associated protein tyrosine phosphatase (BEM)-2	1487	71
1820	gi7141127	Homo sapiens	Ellis-van Creveld syndrome protein	1277	100
1820	gi7271903	Homo sapiens	DWF-1	1271	99
1820	gi13506715	Bos taurus	Ellis-van Creveld syndrome protein	1034	81
1821	AAG72370	Homo sapiens	Human OR-like polypeptide query sequence, SEQ ID NO: 2051.	568	98
1821	AAG71453	Homo sapiens	Human olfactory receptor polypeptide, SEQ ID NO: 1134.	568	98
1821	AAE04556	Homo sapiens	Human G-protein coupled receptor-12 (GCRC-12) protein.	558	100
1822	ABB89189	Homo sapiens	Human polypeptide SEQ ID NO 1565.	388	63
1822	AAM90349	Homo sapiens	Human immune/haematopoietic antigen SEQ ID NO:17942.	126	50
1822	AAB95094	Homo sapiens	Human protein sequence SEQ ID NO:17042.	94	41

Table 3A  
306

SEQ ID	Database entry ID	Description	Results*
912	BL01158	Macrophage migration inhibitory factor family proteins.	BL01158A 21.81 4.971e-37 124-169 BL01158B 17.07 4.343e-23 169-196
912	PF00043	Glutathione S-transferases.	PF00043 21.83 5.333e-14 377-407
913	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 8.200e-17 180-208 PD01719A 12.89 2.364e-13 123-151
913	PR00010	TYPE II EGF-LIKE SIGNATURE	PR00010C 11.16 2.091e-10 817-828
913	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 6.538e-16 812-828 BL01187B 12.04 5.696e-13 569-585 BL01187B 12.04 7.261e-13 696-712 BL01187A 9.98 1.429e-10 508-520 BL01187B 12.04 2.286e-10 484-500 BL01187A 9.98 1.750e-09 796-808
913	BL00022	EGF-like domain proteins.	BL00022B 7.54 1.900e-09 821-828
913	BL01177	Anaphylatoxin domain proteins.	BL01177D 17.50 5.167e-09 503-521
913	BL00281	Bowman-Birk serine protease inhibitors family proteins.	BL00281A 14.18 6.754e-09 479-496
913	BL00799	Granulins proteins.	BL00799B 11.02 7.429e-09 475-511
913	PR00907	THROMBOMODULIN SIGNATURE	PR00907B 11.29 2.479e-11 808-825 PR00907G 11.63 9.660e-10 812-839 PR00907G 11.63 9.745e-10 696-723 PR00907G 11.63 9.027e-09 569-596
914	BL01158	Macrophage migration inhibitory factor family proteins.	BL01158A 21.81 4.073e-33 2-47 BL01158B 17.07 2.884e-20 47-74
915	BL01158	Macrophage migration inhibitory factor family proteins.	BL01158B 17.07 4.343e-23 60-87 BL01158A 21.81 3.656e-22 2-47
916	BL00023	Type II fibronectin collagen-binding domain proteins.	BL00023 24.31 5.091e-27 101-138
916	BL00134	Serine proteases, trypsin family, histidine proteins.	BL00134B 15.99 8.560e-17 626-650 BL00134A 11.96 1.321e-16 469-486 BL00134C 13.45 4.462e-16 662-676
916	BL01253	Type I fibronectin domain proteins.	BL01253C 15.89 9.027e-40 327-366 BL01253B 15.21 5.071e-38 272-316 BL01253H 13.15 7.070e-36 644-679 BL01253E 16.01 1.000e-34 543-580 BL01253F 14.35 2.846e-34 581-620 BL01253A 20.33 7.097e-25 201-230 BL01253G 11.34 8.085e-18 625-639 BL01253D 4.84 7.070e-16 469-483
916	PR00722	CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1) SIGNATURE	PR00722A 12.27 3.793e-14 470-486 PR00722C 10.87 2.059e-13 625-638
916	BL00021	Kringle domain proteins.	BL00021D 24.56 3.000e-32 634-676 BL00021B 13.33 9.217e-17 469-487 BL00021C 22.21 4.293e-12 554-576
916	PR00013	FIBRONECTIN TYPE II REPEAT	PR00013C 12.29 4.273e-19 132-

Table 3A  
307

SEQ ID	Database entry ID	Description	Results*
		SIGNATURE	148 PR00013A 12.26 6.595e-11 105-115 PR00013B 14.75 1.409e-10 116-129
916	BL00495	Apple domain proteins.	BL00495N 11.04 7.987e-21 618-653 BL00495O 13.75 3.311e-17 653-682 BL00495M 8.50 6.243e-10 545-580
916	BL00022	EGF-like domain proteins.	BL00022B 7.54 1.900e-09 267-274
916	PR00018	KRINGLE DOMAIN SIGNATURE	PR00018C 14.30 7.750e-22 331-352 PR00018A 14.52 2.286e-12 286-302 PR00018B 17.75 7.818e-10 302-315 PR00018D 13.51 3.531e-09 356-368
916	PR00010	TYPE II EGF-LIKE SIGNATURE	PR00010A 11.79 4.000e-09 160-172 PR00010C 11.16 8.071e-09 182-193
916	PR00011	TYPE III EGF-LIKE SIGNATURE	PR00011D 14.03 8.826e-09 179-198
916	PD00919	CALCIUM-BINDING PRECURSOR SIGNAL R.	PD00919A 11.53 1.000e-08 164-176
917	BL01248	Laminin-type EGF-like (LE) domain proteins.	BL01248 11.02 4.429e-15 314-327 BL01248 11.02 1.375e-14 379-392 BL01248 11.02 7.975e-11 1011-1024 BL01248 11.02 5.245e-09 1069-1082 BL01248 11.02 5.585e-09 899-912
917	PR00011	TYPE III EGF-LIKE SIGNATURE	PR00011A 14.06 8.579e-19 1060-1079 PR00011D 14.03 9.250e-17 1107-1126 PR00011B 13.08 7.938e-16 846-865 PR00011A 14.06 1.340e-15 535-554 PR00011D 14.03 3.800e-15 846-865 PR00011A 14.06 5.755e-15 846-865 PR00011B 13.08 5.846e-15 1107-1126 PR00011D 14.03 2.286e-14 1060-1079 PR00011B 13.08 2.333e-14 485-504 PR00011B 13.08 6.333e-14 1060-1079 PR00011D 14.03 7.429e-14 485-504 PR00011B 13.08 1.458e-13 535-554 PR00011D 14.03 1.849e-13 535-554 PR00011A 14.06 3.593e-13 1107-1126 PR00011A 14.06 7.254e-13 485-504 PR00011B 13.08 9.847e-13 798-817 PR00011A 14.06 1.581e-12 798-817 PR00011C 24.25 7.623e-12 817-846 PR00011D 14.03 1.148e-11 798-817 PR00011A 14.06 3.492e-11 433-452 PR00011D 14.03 9.262e-11 433-452 PR00011C 24.25 8.800e-10 440-469 PR00011C 24.25 3.143e-09 455-484 PR00011B 13.08 6.548e-09 433-452
917	PD00320	LAMININ CHAIN EGF-LIKE	PD00320A 14.49 4.115e-10 813-

Table 3A  
308

SEQ ID	Database entry ID	Description	Results*
		DOMAIN P.	827 PD00320A 14.49 8.071e-09 1075-1089
918	PR00010	TYPE II EGF-LIKE SIGNATURE	PR00010C 11.16 2.636e-10 132-143
918	BL00790	Receptor tyrosine kinase class V proteins.	BL00790E 29.58 4.621e-10 666-714
918	BL01177	Anaphylatoxin domain proteins.	BL01177E 20.64 6.203e-12 377-404 BL01177D 17.50 3.000e-09 277-295 BL01177C 17.39 4.000e-09 121-140
918	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 8.412e-15 379-395 BL01187B 12.04 5.304e-13 299-315 BL01187B 12.04 7.652e-13 339-355 BL01187B 12.04 8.826e-13 127-143 BL01187B 12.04 6.000e-12 46-62 BL01187B 12.04 5.200e-11 258-274 BL01187B 12.04 6.400e-10 86-102 BL01187A 9.98 8.286e-10 110-122 BL01187A 9.98 5.125e-09 282-294 BL01187A 9.98 6.250e-09 68-80 BL01187A 9.98 8.125e-09 323-335
918	PR00907	THROMBOMODULIN SIGNATURE	PR00907G 11.63 6.774e-10 46-73 PR00907B 11.29 7.545e-10 215-232 PR00907G 11.63 9.321e-10 379-406 PR00907B 11.29 7.805e-09 82-99 PR00907B 11.29 9.232e-09 335-352 PR00907D 5.59 9.427e-09 344-370
919	BL00335	Parathyroid hormone family proteins.	BL00335B 24.23 1.000e-40 133-182 BL00335A 17.98 8.393e-39 80-115
920	BL01177	Anaphylatoxin domain proteins.	BL01177C 17.39 4.508e-10 92-111
920	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 8.043e-13 12-28 BL01187B 12.04 1.000e-11 139-155 BL01187A 9.98 1.474e-11 82-94 BL01187B 12.04 4.900e-11 98-114 BL01187A 9.98 7.429e-10 122-134
920	PR00011	TYPE III EGF-LIKE SIGNATURE	PR00011A 14.06 2.964e-14 288-307 PR00011B 13.08 4.356e-13 288-307 PR00011D 14.03 6.434e-13 376-395 PR00011B 13.08 2.887e-12 376-395 PR00011D 14.03 5.421e-12 288-307 PR00011D 14.03 5.721e-11 331-350 PR00011B 13.08 5.826e-10 331-350 PR00011A 14.06 8.957e-10 376-395
920	DM00864	EGF-LIKE DOMAIN.	DM00864A 15.21 3.314e-09 232-253
920	PR00764	COMPLEMENT C9 SIGNATURE	PR00764F 16.89 5.675e-09 236-257
920	BL00022	EGF-like domain proteins.	BL00022B 7.54 7.300e-09 67-74
920	BL00799	Granulins proteins.	BL00799H 14.15 9.083e-09 227-268

Table 3A  
309

SEQ ID	Database entry ID	Description	Results*
920	PR00907	THROMBOMODULIN SIGNATURE	PR00907B 11.29 8.714e-10 54-71 PR00907B 11.29 9.561e-09 135-152
920	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 1.000e-08 387-430
921	BL00472	Small cytokines (intercrine/chemokine) C-C subfamily signatur.	BL00472B 14.67 2.000e-16 35-53 BL00472A 7.45 4.724e-09 1-13
922	BL00132	Zinc carboxypeptidases, zinc-binding region 1 proteins.	BL00132C 21.35 1.837e-25 227-268 BL00132A 26.07 9.791e-24 149-190 BL00132E 17.72 2.350e-22 297-324 BL00132F 13.26 5.313e-18 325-347 BL00132B 15.93 5.065e-16 197-211 BL00132G 10.94 6.318e-14 382-400 BL00132D 12.70 5.313e-12 271-286
922	PR00765	CARBOXYPEPTIDASE A METALLOPROTEASE (M14) FAMILY SIGNATURE	PR00765B 15.57 2.875e-15 197-212 PR00765D 14.16 8.412e-15 330-344 PR00765C 12.55 2.432e-09 277-286
923	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514C 17.41 3.618e-25 93-130 BL00514E 14.28 8.286e-14 153-170 BL00514D 15.35 2.915e-12 134-147 BL00514G 15.98 4.444e-12 223-253
927	BL00453	FKBP-type peptidyl-prolyl cis-trans isomerase proteins.	BL00453B 23.86 8.500e-20 80-114 BL00453A 15.57 1.000e-15 55-70 BL00453C 9.72 8.650e-11 109-122
937	BL00796	14-3-3 proteins.	BL00796C 17.44 6.250e-38 143-193 BL00796B 10.67 1.514e-36 77-110 BL00796D 17.39 5.696e-34 194-240 BL00796E 14.15 7.353e-29 242-278
937	PR00305	14-3-3 PROTEIN ZETA SIGNATURE	PR00305F 15.95 3.250e-37 248-278 PR00305A 9.33 1.500e-33 77-107 PR00305D 16.34 6.400e-29 194-221 PR00305C 8.68 1.000e-28 159-182 PR00305B 9.99 4.375e-23 126-151 PR00305E 13.01 3.571e-10 221-248
938	BL00303	S-100/ICaBP type calcium binding protein.	BL00303A 21.77 9.526e-31 64-101 BL00303B 26.15 5.737e-30 111-148
938	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 9.471e-11 123-136
940	PD02365	CHAIN FACTOR INTERLEUKIN-12 BETA PRECURSOR IL-1.	PD02365C 7.89 6.680e-10 342-372
940	DM00202	w T-CELL IG HEAVY ALPHA.	DM00202A 9.44 9.813e-09 34-44
944	BL00284	Serpins proteins.	BL00284C 28.56 6.400e-25 319-361 BL00284E 19.15 1.000e-15 508-533 BL00284A 15.64 2.742e-15 200-224 BL00284B 17.99 6.182e-12 292-313 BL00284D 16.34 7.070e-12 430-457
944	PD02080	T-CELL GLYCOPROTEIN CD8 CHAIN SURFACE ALPHA PRE.	PD02080A 10.03 9.750e-10 35-51

Table 3A  
310

SEQ ID	Database entry ID	Description	Results*
944	PR00743	GLYCOSYL HYDROLASE FAMILY 36 SIGNATURE	PR00743B 14.95 8.831e-09 240-261
947	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 4.750e-17 552-565
947	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 2.373e-09 203-257
947	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 4.000e-09 559-568
947	BL00422	Granins proteins.	BL00422E 26.86 8.615e-09 462-498
950	PR00356	TYPE II ANTIFREEZE PROTEIN SIGNATURE	PR00356D 13.09 8.038e-09 174-191
951	BL00615	C-type lectin domain proteins.	BL00615A 16.68 4.316e-13 209-227
951	PR00356	TYPE II ANTIFREEZE PROTEIN SIGNATURE	PR00356B 14.85 6.294e-10 209-227 PR00356A 12.90 7.188e-10 197-210
952	BL00570	Bacterial ring hydroxylating dioxygenases alpha-subunit signa.	BL00570B 19.03 9.357e-09 271-303
953	PD00126	PROTEIN REPEAT DOMAIN TPR NUCLEA.	PD00126A 22.53 1.931e-09 72-93
953	PR00910	LUTEOVIRUS ORF6 PROTEIN SIGNATURE	PR00910A 2.51 3.571e-09 107-120
953	BL00415	Synapsins proteins.	BL00415N 4.29 5.730e-09 13-57
954	PR00901	PHEROMONE B ALPHA-1 RECEPTOR SIGNATURE	PR00901H 14.99 4.706e-09 56-67
958	PR00138	MATRIXIN SIGNATURE	PR00138C 16.41 6.478e-32 178-207 PR00138D 16.56 1.360e-28 236-262 PR00138B 15.82 8.071e-18 155-171 PR00138A 15.14 5.091e-16 108-122 PR00138E 6.01 9.250e-15 269-283
958	BL00024	Hemopexin domain proteins.	BL00024C 22.98 1.000e-40 177-226 BL00024B 21.53 2.636e-33 129-163 BL00024D 17.28 4.086e-31 230-262 BL00024F 11.30 2.731e-22 305-326 BL00024H 11.35 1.947e-14 411-423 BL00024E 7.58 5.500e-14 269-283 BL00024G 13.31 4.115e-13 342-355 BL00024A 11.49 5.050e-13 108-119
958	BL00142	Neutral zinc metalloproteases, zinc-binding region proteins.	BL00142 8.38 3.455e-11 236-247
958	PR00480	ASTACIN FAMILY SIGNATURE	PR00480B 15.41 1.643e-09 231-250
958	BL00546	Matrixins cysteine switch.	BL00546B 20.11 1.000e-40 178-222 BL00546C 16.41 1.771e-31 230-262 BL00546A 19.62 5.304e-28 88-118 BL00546E 10.23 3.323e-22 305-326 BL00546G 16.84 7.300e-19 363-383 BL00546D 10.34 1.486e-14 269-283 BL00546F 12.40 2.800e-13 342-355 BL00546H 10.76 6.625e-12 458-469 BL00546I 10.76

Table 3A

311

SEQ ID	Database entry ID	Description	Results*
			4.512e-09 411-422
959	PD02043	RIBULOSE BISPHOSPHATE CARBOXYLAS.	PD02043A 12.92 5.800e-09 137-171
962	BL00598	Chromo domain proteins.	BL00598 14.45 5.781e-16 63-85
962	PR00504	CHROMODOMAIN SIGNATURE	PR00504C 11.19 5.186e-10 72-85 PR00504B 9.12 3.250e-09 57-72
967	PF00420	NADH-ubiquinone/plastoquinone oxidoreductase chain 4L.	PF00420A 16.63 9.526e-10 89-120
967	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245C 7.84 5.000e-16 185-201 PR00245A 18.03 6.878e-16 59-81 PR00245D 10.47 8.500e-15 221-233 PR00245E 12.40 1.000e-09 238-253
967	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237A 11.48 5.065e-09 26-51
967	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 2.184e-12 90-130 BL00237D 11.23 5.909e-09 229-246
968	BL00059	Zinc-containing alcohol dehydrogenases proteins.	BL00059B 16.08 5.705e-14 223-251
968	BL01162	Quinone oxidoreductase / zeta-crystallin proteins.	BL01162C 22.80 5.846e-14 284-328 BL01162B 18.11 4.205e-11 223-251 BL01162A 15.38 3.805e-09 194-217
970	PF00676	Dehydrogenase E1 component.	PF00676B 24.71 2.800e-39 257-295 PF00676D 14.40 7.545e-24 341-361 PF00676C 16.88 5.737e-23 309-333 PF00676A 12.85 5.050e-12 131-144
970	BL00801	Transketolase proteins.	BL00801D 22.48 7.750e-11 250-290
977	PR00457	ANIMAL HAEM PEROXIDASE SIGNATURE	PR00457E 20.67 9.591e-26 409-436 PR00457D 16.81 5.667e-22 384-405 PR00457G 17.45 1.000e-15 590-611 PR00457B 13.29 4.343e-15 216-232 PR00457C 19.25 3.250e-12 366-385 PR00457A 15.80 5.645e-12 162-174 PR00457H 15.90 1.900e-10 684-699 PR00457F 13.69 6.800e-09 461-472
978	BL00269	Mammalian defensins proteins.	BL00269C 16.52 3.942e-17 95-124 BL00269B 19.17 4.122e-15 57-86 BL00269A 8.53 3.250e-13 31-51
982	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 8.297e-10 79-113
985	PD02283	PROTEIN SPORULATION REPEAT PRECU.	PD02283C 17.54 9.663e-09 78-106
987	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 1.000e-09 247-261 PR00019B 11.36 2.080e-09 316-330 PR00019B 11.36 3.160e-09 244-258 PR00019B 11.36 4.240e-09 100-114 PR00019B 11.36 5.680e-09 148-162
988	BL01215	Mrp family proteins.	BL01215C 18.97 6.447e-36 191-233 BL01215D 30.07 1.000e-33

Table 3A  
312

SEQ ID	Database entry ID	Description	Results*
			250-300 BL01215A 9.75 4.400e-22 101-128 BL01215B 9.34 2.658e-13 134-147
988	PR00364	DISEASE RESISTANCE PROTEIN SIGNATURE	PR00364A 8.19 2.023e-09 105-121
988	BL00746	NifH/frxC family proteins.	BL00746A 24.43 2.819e-09 107-152
988	PR00091	NITROGENASE COMPONENT II SIGNATURE	PR00091A 8.10 4.329e-09 109-123
988	BL01128	Shikimate kinase proteins.	BL01128A 18.84 8.221e-09 106-140
989	PF00094	von Willebrand factor type D domain proteins.	PF00094B 10.43 6.400e-17 491-509
989	PF00054	Laminin G domain proteins.	PF00054B 16.61 7.300e-09 658-670
989	BL00779	Glycoprotein hormones alpha chain proteins.	BL00779A 14.01 7.840e-09 329-356
989	BL01185	C-terminal cystine knot proteins.	BL01185B 21.14 9.153e-09 238-287
992	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 8.986e-11 78-112
992	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 2.000e-09 154-177
994	BL00240	Receptor tyrosine kinase class III proteins.	BL00240B 24.70 4.692e-11 148-172
994	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 5.304e-09 101-111
994	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870D 15.74 9.234e-09 274-309
995	BL01009	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 proteins.	BL01009D 14.19 4.300e-20 127-148 BL01009A 13.75 6.586e-13 57-75 BL01009E 13.50 1.439e-11 159-175
995	PR00837	ALLERGEN V5/TPX-1 FAMILY SIGNATURE	PR00837C 17.21 6.143e-20 126-143 PR00837A 14.77 1.973e-13 57-76 PR00837D 11.12 3.700e-11 160-174
995	PR00838	VENOM ALLERGEN 5 SIGNATURE	PR00838G 16.07 2.033e-17 125-145 PR00838D 8.73 4.214e-09 57-76
996	PD01270	RECEPTOR FC IMMUNOGLOBULIN AFFIN.	PD01270B 22.18 5.567e-18 111-148 PD01270C 19.54 1.167e-17 154-183 PD01270A 17.22 4.960e-14 57-97 PD01270D 24.66 4.284e-09 188-224
999	BL01305	moaA / nifB / pqqE family proteins.	BL01305D 14.97 7.279e-09 7-22
1000	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 4.600e-12 42-65 BL00290B 13.17 1.474e-11 98-116
1002	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 5.846e-15 285-298 PD00066 13.92 1.600e-14 201-214 PD00066 13.92 2.800e-14 313-326 PD00066 13.92 2.000e-13 341-354 PD00066 13.92 5.500e-13 229-242 PD00066 13.92 8.435e-11 257-270
1002	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 9.471e-14 269-286 BL00028 16.07 3.769e-11 241-258



Table 3A  
313

SEQ ID	Database entry ID	Description	Results*
			BL00028 16.07 7.577e-11 325-342 BL00028 16.07 7.577e-11 353-370 BL00028 16.07 6.700e-10 185-202 BL00028 16.07 1.257e-09 297-314
1002	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 1.000e-12 266-280 PR00048A 10.52 3.118e-12 238-252 PR00048A 10.52 8.941e-12 294-308 PR00048B 6.02 2.688e-10 310-320 PR00048B 6.02 3.250e-10 198-208 PR00048A 10.52 3.348e-10 322-336 PR00048A 10.52 3.739e-10 210-224 PR00048A 10.52 5.696e-10 350-364 PR00048A 10.52 1.000e-09 182-196 PR00048B 6.02 1.947e-09 282-292 PR00048B 6.02 5.737e-09 338-348
1003	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 4.706e-14 607-624 BL00028 16.07 6.400e-13 411-428 BL00028 16.07 9.550e-13 579-596 BL00028 16.07 1.000e-12 439-456 BL00028 16.07 1.391e-12 495-512 BL00028 16.07 1.783e-12 523-540 BL00028 16.07 7.652e-12 383-400 BL00028 16.07 9.217e-12 355-372 BL00028 16.07 9.217e-12 663-680 BL00028 16.07 5.846e-11 467-484 BL00028 16.07 6.538e-11 635-652 BL00028 16.07 2.800e-10 327-344 BL00028 16.07 6.143e-09 262-279
1003	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 4.462e-15 399-412 PD00066 13.92 8.615e-15 343-356 PD00066 13.92 1.600e-14 427-440 PD00066 13.92 2.800e-14 511-524 PD00066 13.92 2.800e-14 623-636 PD00066 13.92 5.200e-14 595-608 PD00066 13.92 6.400e-14 567-580 PD00066 13.92 8.800e-14 371-384 PD00066 13.92 7.000e-13 315-328 PD00066 13.92 9.000e-13 483-496 PD00066 13.92 6.870e-11 455-468 PD00066 13.92 7.600e-09 539-552 PD00066 13.92 7.900e-09 651-664
1003	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 9.182e-15 492-506 PR00048A 10.52 1.750e-14 408-422 PR00048A 10.52 7.000e-14 576-590 PR00048A 10.52 1.000e-13 604-618 PR00048A 10.52 3.571e-13 380-394 PR00048A 10.52 4.214e-13 464-478 PR00048A 10.52 5.235e-12 352-366 PR00048B 6.02 1.000e-11 424-434 PR00048B 6.02 1.692e-11 340-350 PR00048B 6.02 7.231e-11 536-546 PR00048B 6.02 7.231e-11 620-630 PR00048A 10.52 8.579e-

Table 3A  
314

SEQ ID	Database entry ID	Description	Results*
			11 259-273 PR00048A 10.52 9.053e-11 436-450 PR00048B 6.02 9.308e-11 508-518 PR00048B 6.02 1.563e-10 592-602 PR00048B 6.02 2.125e-10 396-406 PR00048A 10.52 4.130e-10 660-674 PR00048A 10.52 4.522e-10 520- 534 PR00048A 10.52 6.478e-10 324-338 PR00048B 6.02 1.474e-09 564-574 PR00048B 6.02 2.421e-09 676-686 PR00048A 10.52 2.800e- 09 548-562 PR00048A 10.52 2.800e-09 632-646 PR00048B 6.02 2.895e-09 368-378 PR00048B 6.02 1.000e-08 648-658
1007	BL00478	LIM domain proteins.	BL00478B 14.79 3.739e-14 469- 484 BL00478B 14.79 3.500e-12 411-426 BL00478B 14.79 6.000e- 12 536-551
1007	DM00984	w MYOD MYOBLAST DETERMINATION SHORT.	DM00984B 15.18 4.822e-25 426- 481 DM00984C 7.66 8.036e-11 481-495
1007	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115Z 3.12 7.022e-09 260-309
1008	DM00475	w LOW TRANSPOSASE SAPA 12K.	DM00475B 12.12 8.269e-09 115- 135
1011	PR00456	RIBOSOMAL PROTEIN P2 SIGNATURE	PR00456F 5.86 6.400e-12 76-88
1012	PF00756	Putative esterase.	PF00756C 14.12 7.692e-10 103- 133
1016	PR00310	ANTI-PROLIFERATIVE PROTEIN BTG1 FAMILY SIGNATURE	PR00310D 9.10 1.540e-37 183-213 PR00310C 12.74 5.286e-35 153- 183 PR00310A 11.17 7.000e-27 16-41 PR00310E 13.58 6.914e-24 229-249 PR00310B 10.59 3.687e- 23 123-153
1016	BL00960	BTG1 family proteins.	BL00960B 24.47 3.288e-26 116- 161 BL00960C 12.68 3.647e-26 180-202 BL00960A 10.98 5.304e- 12 14-26
1017	PD01066	PROTEIN ZINC FINGER ZINC- FINGER METAL-BINDING NU.	PD01066 19.43 6.143e-32 44-83
1017	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 6.063e-15 523-540 BL00028 16.07 7.188e-15 383-400 BL00028 16.07 3.700e-13 467-484 BL00028 16.07 5.950e-13 439-456 BL00028 16.07 8.650e-13 271-288 BL00028 16.07 4.115e-11 355-372 BL00028 16.07 6.885e-11 327-344 BL00028 16.07 2.800e-10 411-428 BL00028 16.07 3.100e-10 495-512
1017	PD00066	PROTEIN ZINC-FINGER METAL- BINDI.	PD00066 13.92 6.625e-18 399-412 PD00066 13.92 2.385e-15 371-384 PD00066 13.92 8.615e-15 343-356 PD00066 13.92 3.500e-13 511-524 PD00066 13.92 1.000e-12 483-496

Table 3A  
315

SEQ ID	Database entry ID	Description	Results*
			PD00066 13.92 1.000e-11 427-440 PD00066 13.92 3.769e-10 455-468 PD00066 13.92 5.800e-09 539-552
1017	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 6.727e-15 520-534 PR00048A 10.52 1.750e-14 380-394 PR00048A 10.52 3.250e-14 408-422 PR00048A 10.52 4.857e-13 464-478 PR00048A 10.52 1.529e-12 436-450 PR00048A 10.52 2.059e-12 352-366 PR00048B 6.02 3.769e-11 368-378 PR00048B 6.02 5.846e-11 396-406 PR00048B 6.02 5.846e-11 508-518 PR00048A 10.52 1.783e-10 492-506 PR00048B 6.02 8.313e-10 480-490 PR00048A 10.52 4.240e-09 324-338 PR00048B 6.02 5.737e-09 340-350 PR00048A 10.52 8.200e-09 146-160
1020	BL00478	LIM domain proteins.	BL00478B 14.79 6.000e-15 219-234 BL00478B 14.79 8.250e-12 99-114 BL00478B 14.79 8.250e-12 160-175 BL00478B 14.79 2.800e-11 282-297
1024	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 9.400e-38 17-56
1024	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 1.600e-14 230-243 PD00066 13.92 8.200e-14 286-299 PD00066 13.92 8.200e-14 398-411 PD00066 13.92 1.000e-13 342-355 PD00066 13.92 6.478e-11 314-327 PD00066 13.92 6.478e-11 370-383 PD00066 13.92 6.870e-11 258-271
1024	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 6.850e-13 354-371 BL00028 16.07 4.522e-12 242-259 BL00028 16.07 8.826e-12 214-231 BL00028 16.07 9.609e-12 270-287 BL00028 16.07 2.731e-11 326-343 BL00028 16.07 1.900e-10 382-399 BL00028 16.07 8.500e-10 298-315 BL00028 16.07 5.629e-09 410-427
1024	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479A 19.86 7.319e-09 201-224
1024	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 4.273e-15 351-365 PR00048A 10.52 5.500e-13 239-253 PR00048A 10.52 5.500e-13 267-281 PR00048A 10.52 8.941e-12 323-337 PR00048A 10.52 5.263e-11 379-393 PR00048A 10.52 1.783e-10 407-421 PR00048B 6.02 3.250e-10 339-349 PR00048B 6.02 8.875e-10 227-237 PR00048B 6.02 1.000e-09 283-293 PR00048B 6.02 1.000e-09 395-405 PR00048A 10.52 1.360e-

Table 3A  
316

SEQ ID	Database entry ID	Description	Results*
			09 295-309 PR00048B 6.02 2.895e-09 423-433 PR00048A 10.52 7.480e-09 211-225
1025	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824B 9.21 4.892e-09 394-414
1025	PR00966	POTYVIRUS NUCLEAR INCLUSION A CYSTEINE PROTEASE (C4) SIGNATURE	PR00966H 13.17 9.727e-09 215-235
1026	BL01282	BIR repeat proteins.	BL01282B 30.49 3.829e-09 12-51
1026	PF00992	Troponin.	PF00992A 16.67 5.263e-09 136-171
1026	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 8.333e-09 31-40
1029	PD02283	PROTEIN SPORULATION REPEAT PRECU.	PD02283C 17.54 5.613e-09 171-199
1029	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 5.213e-10 91-137 BL00203 13.94 5.041e-09 71-117 BL00203 13.94 8.898e-09 53-99
1031	BL00170	Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatur.	BL00170B 20.97 1.000e-40 48-88 BL00170C 18.49 1.000e-40 95-140 BL00170A 17.08 2.452e-16 18-45
1031	PR00153	CYCLOPHILIN PEPTIDYL-PROLYL CIS-TRANS ISOMERASE SIGNATURE	PR00153C 11.01 4.375e-20 96-112 PR00153B 11.57 1.500e-17 53-66 PR00153E 9.10 7.632e-17 124-140 PR00153D 11.99 6.400e-16 111-124 PR00153A 12.98 3.093e-11 24-40
1031	PF00638	RanBP1 domain proteins.	PF00638 11.91 5.569e-09 68-83
1035	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.149e-29 6-45
1035	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 1.692e-15 234-247 PD00066 13.92 3.077e-15 374-387 PD00066 13.92 3.077e-15 458-471 PD00066 13.92 5.846e-15 486-499 PD00066 13.92 3.400e-14 150-163 PD00066 13.92 5.800e-14 206-219 PD00066 13.92 5.800e-14 290-303 PD00066 13.92 5.800e-14 318-331 PD00066 13.92 8.800e-14 514-527 PD00066 13.92 3.500e-13 346-359 PD00066 13.92 7.500e-13 402-415 PD00066 13.92 8.714e-12 178-191 PD00066 13.92 3.700e-09 430-443 PD00066 13.92 5.500e-09 262-275
1035	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 4.750e-14 215-229 PR00048A 10.52 4.750e-14 467-481 PR00048A 10.52 7.000e-14 355-369 PR00048A 10.52 4.857e-13 299-313 PR00048A 10.52 5.500e-13 99-113 PR00048A 10.52 5.500e-13 327-341 PR00048A 10.52 1.000e-12 495-509 PR00048A 10.52 2.059e-12 271-285 PR00048A 10.52 2.588e-12 523-537 PR00048A 10.52 8.412e-12 243-257 PR00048A

Table 3A  
317

SEQ ID	Database entry ID	Description	Results*
			10.52 4.316e-11 131-145 PR00048B 6.02 4.462e-11 115-125 PR00048A 10.52 8.579e-11 439-453 PR00048A 10.52 9.053e-11 383-397 PR00048A 10.52 1.391e-10 159-173 PR00048B 6.02 1.563e-10 455-465 PR00048A 10.52 3.739e-10 411-425 PR00048B 6.02 6.625e-10 203-213 PR00048B 6.02 6.625e-10 287-297 PR00048A 10.52 3.160e-09 551-565 PR00048A 10.52 4.600e-09 187-201 PR00048B 6.02 8.579e-09 147-157
1035	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 1.900e-13 358-375 BL00028 16.07 2.350e-13 102-119 BL00028 16.07 8.826e-12 470-487 BL00028 16.07 2.038e-11 554-571 BL00028 16.07 5.500e-11 190-207 BL00028 16.07 5.846e-11 274-291 BL00028 16.07 6.192e-11 218-235 BL00028 16.07 8.269e-11 330-347 BL00028 16.07 1.300e-10 526-543 BL00028 16.07 1.600e-10 414-431 BL00028 16.07 3.700e-10 386-403 BL00028 16.07 6.700e-10 442-459 BL00028 16.07 7.000e-10 302-319 BL00028 16.07 1.000e-09 134-151 BL00028 16.07 3.314e-09 246-263 BL00028 16.07 1.000e-08 498-515
1038	BL01130	Sulfate transporters proteins.	BL01130A 21.63 7.407e-25 331-385 BL01130B 23.34 2.286e-23 429-481
1038	DM01292	ESICULAR LUMEN DOMAIN.	DM01292I 12.82 9.400e-10 148-190 DM01292I 12.82 9.400e-10 591-633
1042	PD02327	GLYCOPROTEIN ANTIGEN PRECURSOR IMMUNOGLO.	PD02327B 19.84 5.574e-10 169-191
1042	PD02365	CHAIN FACTOR INTERLEUKIN-12 BETA PRECURSOR IL-1.	PD02365C 7.89 4.196e-09 365-395
1043	BL00983	Ly-6 / u-PAR domain proteins.	BL00983C 12.69 3.500e-15 142-158 BL00983B 8.19 1.643e-12 84-94 BL00983A 5.84 7.261e-10 74-83
1047	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290B 13.17 6.400e-22 281-299 BL00290A 20.89 4.600e-16 34-57 BL00290A 20.89 2.080e-10 224-247
1051	BL01221	PMP-22 / EMP / MP20 family proteins.	BL01221B 13.29 6.745e-09 54-68
1052	BL00983	Ly-6 / u-PAR domain proteins.	BL00983C 12.69 4.981e-09 46-62
1052	BL00272	Snake toxins proteins.	BL00272C 8.27 8.326e-09 50-62
1054	BL00420	Speract receptor repeat proteins domain proteins.	BL00420B 22.67 4.627e-30 723-778 BL00420C 11.90 9.100e-13 809-820
1054	PR00258	SPERACT RECEPTOR SIGNATURE	PR00258B 9.63 3.813e-15 738-750 PR00258E 13.33 2.047e-12 808-

Table 3A  
318

SEQ ID	Database entry ID	Description	Results*
			821 PR00258C 9.05 2.837e-10 753-764
1054	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514G 15.98 4.326e-09 542-572
1055	BL01212	ATP P2X receptors proteins.	BL01212A 34.89 1.000e-40 41-94 BL01212E 24.87 1.000e-40 225-280 BL01212G 11.86 3.700e-34 309-337 BL01212D 11.42 9.609e-27 182-206 BL01212B 19.25 8.393e-21 126-151 BL01212F 10.12 2.421e-15 290-301 BL01212C 8.40 2.500e-14 158-169
1056	PR00920	SPUMAVIRUS ASPARTIC PROTEASE (A9) SIGNATURE	PR00920C 13.24 7.310e-09 149-171
1057	BL00682	ZP domain proteins.	BL00682C 20.71 1.706e-12 439-464
1057	BL00025	P-type 'Trefoil' domain proteins.	BL00025 17.17 5.645e-09 231-252
1059	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 4.273e-14 98-138
1059	PR00854	PROSTAGLANDIN D RECEPTOR SIGNATURE	PR00854E 10.50 4.649e-26 236-260 PR00854B 7.30 8.154e-21 41-59 PR00854G 10.66 1.783e-18 341-358 PR00854D 9.41 2.500e-18 185-201 PR00854A 15.24 9.077e-18 6-21 PR00854H 14.71 6.203e-17 369-390 PR00854C 12.92 1.643e-12 93-105 PR00854F 12.83 9.682e-11 321-333
1059	PR00856	PROSTACYCLIN (PROSTANOID IP) RECEPTOR SIGNATURE	PR00856E 9.82 1.724e-09 178-195
1060	BL01271	Sodium:sulfate symporter family proteins.	BL01271D 25.26 1.000e-40 480-535 BL01271B 12.02 6.400e-24 208-233 BL01271A 8.06 7.955e-23 132-152 BL01271C 13.62 7.429e-20 407-429
1062	PF00798	Arenavirus glycoprotein.	PF00798I 18.55 8.811e-09 53-90
1064	BL01017	Ergosterol biosynthesis ERG4/ERG24 family proteins.	BL01017D 20.82 1.000e-40 232-278 BL01017F 23.34 9.196e-35 291-344 BL01017C 15.91 7.324e-23 181-207 BL01017B 12.69 9.419e-17 166-181
1065	BL00874	Bacterial type II secretion system protein F proteins.	BL00874B 29.89 9.724e-09 414-469
1066	PD01270	RECEPTOR FC IMMUNOGLOBULIN AFFIN.	PD01270C 19.54 2.895e-16 43-72
1066	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 8.435e-09 183-193
1067	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 3.455e-14 77-117
1067	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237C 15.69 1.257e-10 91-114 PR00237E 13.03 9.100e-10 175-199
1067	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 9.581e-18 46-68 PR00245C 7.84 4.780e-13 214-230 PR00245E 12.40 6.741e-09 267-282
1067	PR00534	MELANOCORTIN RECEPTOR FAMILY SIGNATURE	PR00534A 11.49 9.229e-09 38-51

Table 3A  
319

SEQ ID	Database entry ID	Description	Results*
1069	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 5.909e-15 409-440 BL00107B 13.31 4.214e-11 484-500
1069	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109E 14.41 4.353e-09 549-572
1071	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 4.789e-13 222-245
1072	PR00308	TYPE I ANTIFREEZE PROTEIN SIGNATURE	PR00308C 3.83 8.892e-10 4-14 PR00308C 3.83 8.892e-10 5-15 PR00308C 3.83 8.013e-09 3-13
1072	PR00698	C.ELEGANS SRG FAMILY INTEGRAL MEMBRANE PROTEIN SIGNATURE	PR00698E 14.43 8.714e-09 111-137
1075	PF00023	Ank repeat proteins.	PF00023A 16.03 7.000e-11 69-85 PF00023B 14.20 2.636e-09 131-141
1075	PD00078	REPEAT PROTEIN ANK NUCLEAR ANKYR.	PD00078B 13.14 6.087e-09 128-141
1075	PR00806	VINCULIN SIGNATURE	PR00806C 11.07 8.839e-09 350-368
1075	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 9.505e-09 135-190 PF00791B 28.49 9.835e-09 69-124
1076	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 5.610e-11 174-214 BL00237C 13.19 4.176e-10 317-344
1076	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237F 13.57 7.677e-11 322-347 PR00237E 13.03 6.100e-10 276-300 PR00237A 11.48 8.839e-09 103-128
1077	BL00216	Sugar transport proteins.	BL00216B 27.64 1.831e-09 139-189
1079	BL00462	Gamma-glutamyltranspeptidase proteins.	BL00462A 20.89 4.000e-20 108-151 BL00462D 23.07 7.256e-12 356-396 BL00462B 17.88 9.153e-12 183-220
1080	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115Z 3.12 2.125e-09 1363-1412 BL00115Z 3.12 6.096e-09 1349-1398
1083	BL00272	Snake toxins proteins.	BL00272C 8.27 9.182e-10 109-121
1083	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.348e-14 366-397 BL00107B 13.31 4.176e-09 441-457
1083	PR00653	ACTIVIN TYPE II RECEPTOR SIGNATURE	PR00653D 13.25 7.200e-09 385-407
1083	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109E 14.41 6.727e-11 535-558 PR00109D 17.04 7.609e-09 442-465
1083	BL00983	Ly-6 / u-PAR domain proteins.	BL00983C 12.69 9.135e-09 105-121
1084	BL00272	Snake toxins proteins.	BL00272C 8.27 9.182e-10 109-121
1084	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.348e-14 428-459 BL00107B 13.31 4.176e-09 503-519
1084	PR00653	ACTIVIN TYPE II RECEPTOR SIGNATURE	PR00653D 13.25 7.200e-09 447-469

Table 3A  
320

SEQ ID	Database entry ID	Description	Results*
1084	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109E 14.41 6.727e-11 597-620 PR00109D 17.04 7.609e-09 504-527
1084	BL00983	Ly-6 / u-PAR domain proteins.	BL00983C 12.69 9.135e-09 105-121
1085	PR00541	MUSCARINIC M4 RECEPTOR SIGNATURE	PR00541C 8.06 7.726e-09 486-507
1086	PR00541	MUSCARINIC M4 RECEPTOR SIGNATURE	PR00541C 8.06 7.726e-09 755-776
1087	PD01101	INHIBITOR HEAVY CHAIN CHANNEL IN.	PD01101B 21.53 3.318e-22 343-396
1088	BL00129	Glycosyl hydrolases family 31 proteins.	BL00129A 26.21 2.400e-28 114-160 BL00129D 16.76 6.806e-26 364-408 BL00129C 15.12 5.295e-24 326-354 BL00129E 22.60 4.857e-23 428-464 BL00129B 19.19 4.436e-15 225-252 BL00129F 26.19 2.500e-13 544-582
1090	PR00887	STRUCTURE-SPECIFIC RECOGNITION PROTEIN SIGNATURE	PR00887A 11.39 1.643e-22 343-360 PR00887F 12.74 2.000e-22 498-516 PR00887B 9.94 3.250e-22 365-382 PR00887C 13.16 4.000e-22 388-405 PR00887E 10.36 5.200e-22 480-499 PR00887H 11.84 8.313e-22 537-556 PR00887G 14.17 9.438e-20 521-538 PR00887D 15.12 8.313e-17 453-467
1090	PR00886	HIGH MOBILITY GROUP (HMG1/HMG2) PROTEIN SIGNATURE	PR00886C 11.84 8.500e-13 696-715 PR00886A 10.08 3.192e-10 710-733
1090	PD02448	TRANSCRIPTION PROTEIN DNA-BINDIN.	PD02448A 9.37 5.576e-10 686-725
1090	BL00353	HMG1/2 proteins.	BL00353B 11.47 8.244e-24 664-714 BL00353A 9.60 2.549e-09 674-723
1091	BL00284	Serpins proteins.	BL00284C 28.56 4.000e-25 472-514 BL00284D 16.34 5.655e-17 578-605 BL00284A 15.64 2.742e-15 341-365 BL00284E 19.15 4.818e-15 659-684 BL00284B 17.99 3.667e-14 445-466 BL00284A 15.64 2.600e-11 375-399
1092	PR00839	V8 SERINE PROTEASE FAMILY SIGNATURE	PR00839B 11.20 8.119e-10 357-375
1097	BL00605	ATP synthase c subunit proteins.	BL00605 27.67 3.172e-33 79-133
1097	PR00124	ATP SYNTHASE C SUBUNIT SIGNATURE	PR00124C 12.42 6.400e-18 113-139 PR00124A 8.81 8.054e-14 75-95 PR00124B 14.66 6.897e-12 96-112
1098	BL00450	Aconitase family proteins.	BL00450B 42.34 8.393e-30 386-441 BL00450D 21.14 2.800e-18 665-689 BL00450E 16.34 8.875e-13 710-725 BL00450B 42.34



Table 3A  
321

SEQ ID	Database entry ID	Description	Results*
			6.400e-12 446-501 BL00450A 13.76 2.406e-11 351-365 BL00450C 11.95 6.657e-10 612-622
1098	PR00415	ACONITASE FAMILY SIGNATURE	PR00415D 12.72 5.696e-16 390-406 PR00415I 13.62 4.115e-15 675-689 PR00415G 14.24 8.105e-15 548-563 PR00415C 13.34 7.828e-14 376-390 PR00415E 10.04 7.828e-14 452-466 PR00415F 11.66 7.273e-13 466-480 PR00415H 12.39 9.700e-13 613-625 PR00415A 11.15 1.621e-10 323-337 PR00415B 8.14 9.036e-09 347-356
1104	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 1.000e-11 139-149
1104	BL00240	Receptor tyrosine kinase class III proteins.	BL00240B 24.70 4.255e-09 85-109
1106	BL01002	Translationally controlled tumor protein.	BL01002C 21.97 6.143e-26 79-110 BL01002A 13.19 1.360e-24 1-24 BL01002B 7.39 3.118e-14 48-62
1107	PF00997	Kappa casein.	PF00997D 9.95 8.306e-09 513-548
1109	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 1.391e-09 42-55
1110	BL01310	ATP1G1 / PLM / MAT8 family proteins.	BL01310 14.74 8.981e-24 99-135
1112	PR00764	COMPLEMENT C9 SIGNATURE	PR00764B 13.56 2.250e-11 122-143
1112	PR00261	LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR SIGNATURE	PR00261E 11.08 6.308e-09 127-149 PR00261F 11.57 7.152e-09 127-149
1115	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 4.273e-14 188-228
1115	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 3.250e-19 157-179 PR00245B 10.38 1.918e-09 275-290
1115	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237C 15.69 4.150e-09 202-225
1116	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 2.658e-12 163-203
1116	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 9.325e-19 132-154 PR00245C 7.84 4.073e-15 311-327 PR00245B 10.38 5.500e-13 250-265 PR00245E 12.40 7.618e-13 364-379 PR00245D 10.47 4.673e-09 347-359
1116	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237C 15.69 6.400e-10 177-200 PR00237G 19.63 5.814e-09 345-372
1119	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 7.955e-13 969-997 PD01719A 12.89 8.111e-09 305-333
1120	BL00355	HMG14 and HMG17 proteins.	BL00355 5.97 1.692e-37 18-49
1120	PR00925	NONHISTONE CHROMOSOMAL PROTEIN HMG17 FAMILY	PR00925A 5.47 2.800e-19 18-33 PR00925B 3.73 3.400e-16 34-47

Table 3A  
322

SEQ ID	Database entry ID	Description	Results*
		SIGNATURE	PR00925D 6.56 2.200e-13 66-77 PR00925C 5.57 8.235e-09 47-58
1124	BL00615	C-type lectin domain proteins.	BL00615A 16.68 4.240e-11 210-228
1125	PD02327	GLYCOPROTEIN ANTIGEN PRECURSOR IMMUNOGLO.	PD02327B 19.84 2.091e-09 191-213
1125	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 7.652e-09 338-348
1127	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870D 15.74 8.755e-09 96-131
1133	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 8.364e-14 122-144 PR00245C 7.84 9.280e-13 300-316 PR00245B 10.38 4.600e-11 240-255 PR00245E 12.40 7.623e-10 353-368
1133	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 5.371e-13 153-193 BL00237D 11.23 7.750e-10 344-361
1133	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237G 19.63 6.063e-12 334-361 PR00237C 15.69 6.175e-09 167-190
1134	BL00221	MIP family proteins.	BL00221B 10.22 1.871e-11 141-152 BL00221D 12.33 2.174e-11 240-255 BL00221E 8.47 9.710e-11 307-318 BL00221A 6.39 5.935e-09 92-103
1134	PR00783	MAJOR INTRINSIC PROTEIN FAMILY SIGNATURE	PR00783B 15.98 4.130e-15 127-152 PR00783F 12.33 9.156e-14 308-329 PR00783A 12.72 7.462e-12 88-108 PR00783E 16.78 8.263e-10 128-151 PR00783C 13.54 1.340e-09 164-184 PR00783E 16.78 6.754e-09 226-249
1136	PD02886	GLYCOPROTEIN PRECURSOR IMMUNOGLOBULIN FOL.	PD02886C 21.92 7.907e-10 112-151
1137	PR00122	VACUOLAR ATP SYNTHASE 16 KD SUBUNIT SIGNATURE	PR00122C 8.20 1.000e-33 104-131 PR00122B 8.60 2.125e-28 56-81 PR00122D 9.97 4.375e-28 131-155 PR00122A 11.44 6.053e-19 30-55
1137	BL00605	ATP synthase c subunit proteins.	BL00605 27.67 1.778e-10 94-148
1137	PR00124	ATP SYNTHASE C SUBUNIT SIGNATURE	PR00124C 12.42 2.161e-10 128-154
1138	BL00665	Dihydrodipicolinate synthetase proteins.	BL00665B 30.33 8.265e-12 52-105 BL00665D 14.76 1.000e-11 164-187 BL00665C 25.58 5.832e-11 105-156
1138	PR00146	DIHYDRODIPICOLINATE SYNTHASE SIGNATURE	PR00146D 16.26 2.525e-10 163-181
1139	BL00456	Sodium:solute symporter family proteins.	BL00456C 24.55 4.886e-28 165-220 BL00456A 22.59 3.127e-27 27-82 BL00456B 18.94 1.220e-17 103-133
1139	BL00415	Synapsins proteins.	BL00415O 3.44 6.270e-09 514-552
1139	BL00136	Serine proteases, subtilase family, aspartic acid proteins.	BL00136B 9.63 7.796e-09 773-786

Table 3A  
323

SEQ ID	Database entry ID	Description	Results*
1141	BL00310	Lysosome-associated membrane glycoproteins duplicated domain proteins.	BL00310F 23.26 4.162e-09 194-249
1143	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 4.115e-18 126-166 BL00237C 13.19 7.545e-15 263-290 BL00237D 11.23 8.962e-11 324-341
1143	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237G 19.63 7.120e-15 314-341 PR00237F 13.57 2.565e-14 268-293 PR00237C 15.69 6.667e-12 140-163 PR00237A 11.48 8.125e-11 63-88 PR00237B 13.50 1.563e-10 96-118 PR00237E 13.03 3.118e-09 226-250
1144	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 2.068e-09 7-22
1145	BL01159	WW/rsp5/WWP domain proteins.	BL01159 13.85 1.310e-14 387-402
1145	PR00403	WW DOMAIN SIGNATURE	PR00403B 12.19 6.906e-15 387-402 PR00403A 16.82 5.200e-11 373-387
1145	BL01179	Phosphotyrosine interaction domain proteins (PID) profile.	BL01179A 12.63 8.286e-11 394-406 BL01179B 15.18 7.968e-10 667-682
1147	BL00594	Aromatic amino acids permeases proteins.	BL00594A 16.75 3.851e-09 107-151
1148	PR00399	SYNAPTOTAGMIN SIGNATURE	PR00399B 14.27 1.305e-09 242-256
1148	PR00360	C2 DOMAIN SIGNATURE	PR00360B 13.61 6.318e-09 279-293
1150	BL00291	Prion protein.	BL00291A 4.49 8.241e-09 21-56
1152	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 6.447e-12 210-250
1152	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237G 19.63 3.512e-09 392-419 PR00237C 15.69 4.825e-09 224-247
1152	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 1.500e-20 179-201 PR00245B 10.38 3.571e-16 297-312 PR00245E 12.40 1.000e-12 411-426 PR00245D 10.47 1.000e-10 394-406 PR00245C 7.84 6.727e-09 358-374
1153	PR00962	LETHAL(2) GIANT LARVAE PROTEIN SIGNATURE	PR00962B 11.98 2.800e-28 310-333 PR00962G 15.71 5.655e-28 609-634 PR00962D 10.40 1.225e-27 451-475 PR00962F 12.39 6.786e-23 568-588 PR00962H 13.32 9.710e-23 639-659 PR00962I 11.68 3.829e-22 708-728 PR00962C 8.00 4.250e-22 362-383 PR00962A 13.28 7.612e-22 17-36 PR00962E 8.81 1.628e-20 531-550
1153	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320A 16.74 9.122e-09 454-469
1155	BL00218	Amino acid permeases proteins.	BL00218D 21.49 2.038e-10 385-430 BL00218E 23.30 6.400e-10 466-506 BL00218B 21.44 5.790e-

Table 3A  
324

SEQ ID	Database entry ID	Description	Results*
			09 217-249
1155	BL00341	Surfactant associated polypeptide SP-C palmitoylation site proteins.	BL00341B 8.70 7.895e-09 54-88
1158	PR00783	MAJOR INTRINSIC PROTEIN FAMILY SIGNATURE	PR00783C 13.54 1.474e-17 31-51
1158	BL00221	MIP family proteins.	BL00221B 10.22 1.643e-14 8-19
1158	PD00302	PROTEASE POLYPROTEIN HYDROLASE ASP.	PD00302B 9.52 1.360e-14 261-277 PD00302A 6.33 3.323e-11 198-209
1158	PF00692	dUTPase.	PF00692B 8.14 3.613e-11 113-124
1158	DM00892	3 RETROVIRAL PROTEINASE.	DM00892C 23.55 4.818e-13 292-326 DM00892B 9.78 1.000e-08 264-270
1160	PD00320	LAMININ CHAIN EGF-LIKE DOMAIN P.	PD00320A 14.49 8.286e-12 1681-1695
1160	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 1.225e-09 1745-1761
1160	BL01248	Laminin-type EGF-like (LE) domain proteins.	BL01248 11.02 9.308e-18 1705-1718 BL01248 11.02 4.000e-15 393-406 BL01248 11.02 1.900e-11 321-334 BL01248 11.02 6.094e-09 445-458
1160	PR00877	PLANT PEC FAMILY METALLOTHIONEIN SIGNATURE	PR00877D 4.18 6.516e-09 1308-1316
1160	PF00054	Laminin G domain proteins.	PF00054B 16.61 8.200e-09 558-570
1160	PR00011	TYPE III EGF-LIKE SIGNATURE	PR00011B 13.08 2.841e-17 547-566 PR00011D 14.03 4.150e-16 547-566 PR00011B 13.08 7.750e-16 1370-1389 PR00011A 14.06 1.170e-15 1277-1296 PR00011D 14.03 4.000e-15 501-520 PR00011A 14.06 4.736e-15 547-566 PR00011A 14.06 1.327e-14 1370-1389 PR00011A 14.06 5.909e-14 1416-1435 PR00011D 14.03 7.980e-14 1370-1389 PR00011D 14.03 3.717e-13 1416-1435 PR00011D 14.03 6.434e-13 1277-1296 PR00011B 13.08 8.932e-13 1277-1296 PR00011A 14.06 2.161e-12 436-455 PR00011D 14.03 6.053e-12 436-455 PR00011B 13.08 1.138e-11 436-455 PR00011B 13.08 6.954e-11 1416-1435 PR00011A 14.06 8.062e-11 1318-1337 PR00011A 14.06 1.913e-10 695-714 PR00011B 13.08 2.043e-10 644-663 PR00011D 14.03 3.215e-10 695-714 PR00011C 24.25 4.600e-10 651-680 PR00011C 24.25 9.100e-10 1423-1452 PR00011D 14.03 1.261e-09 644-663 PR00011B 13.08 1.370e-09 695-714 PR00011A 14.06 1.986e-09 644-663 PR00011B 13.08 3.096e-

Table 3A  
325

SEQ ID	Database entry ID	Description	Results*
			09 1318-1337 PR00011D 14.03 8.435e-09 591-610 PR00011C 24.25 9.857e-09 1769-1798
1161	PD00919	CALCIUM-BINDING PRECURSOR SIGNAL R.	PD00919A 11.53 8.377e-10 185- 197
1161	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 9.486e-10 158- 174 BL01187B 12.04 2.800e-09 196-212
1161	PR00011	TYPE III EGF-LIKE SIGNATURE	PR00011D 14.03 4.158e-12 39-58 PR00011B 13.08 2.973e-09 39-58
1161	PR00010	TYPE II EGF-LIKE SIGNATURE	PR00010C 11.16 5.929e-09 163- 174
1161	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243H 17.53 7.276e-09 65-91
1161	BL01185	C-terminal cystine knot proteins.	BL01185B 21.14 9.047e-09 137- 186
1161	DM00060	338 kw NEUREXIN ALPHA III CYSTEINE.	DM00060 6.92 9.460e-09 108-118
1161	BL01248	Laminin-type EGF-like (LE) domain proteins.	BL01248 11.02 9.660e-09 48-61
1162	BL00252	Interferon alpha, beta and delta family proteins.	BL00252A 18.49 6.657e-23 35-72 BL00252B 19.78 9.125e-16 73-124
1162	PR00266	INTERFERON ALPHA AND BETA SUBUNIT SIGNATURE	PR00266A 13.61 1.000e-13 67-80
1163	PR00264	INTERLEUKIN-1 SIGNATURE	PR00264B 20.98 8.453e-11 63-90 PR00264C 17.77 1.851e-10 103- 132
1164	BL00249	Platelet-derived growth factor (PDGF) family proteins.	BL00249 18.21 3.250e-34 75-105
1165	BL00253	Interleukin-1 proteins.	BL00253D 25.67 3.464e-11 95-135
1165	PR00264	INTERLEUKIN-1 SIGNATURE	PR00264C 17.77 3.294e-17 95-124 PR00264B 20.98 6.250e-09 56-83
1166	PR00573	INTERLEUKIN 8B RECEPTOR SIGNATURE	PR00573D 15.57 1.450e-20 114- 130
1166	PR00427	INTERLEUKIN-8 RECEPTOR SIGNATURE	PR00427E 7.08 9.446e-18 6-22 PR00427G 10.85 2.667e-16 98-115 PR00427F 11.77 9.500e-13 48-65
1166	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237F 13.57 4.789e-16 23-48 PR00237G 19.63 1.844e-12 70-97
1166	BL00237	G-protein coupled receptors proteins.	BL00237C 13.19 1.857e-16 18-45 BL00237D 11.23 6.464e-10 80-97
1167	PR00414	PALMITOYL PROTEIN THIOESTERASE SIGNATURE	PR00414A 10.74 1.000e-31 26-49 PR00414D 17.19 1.450e-27 151- 171 PR00414E 16.76 1.600e-27 185-205 PR00414G 12.38 4.000e- 25 258-277 PR00414B 12.48 4.214e-25 66-86 PR00414C 13.14 1.000e-22 106-123
1168	BL00284	Serpins proteins.	BL00284C 28.56 4.000e-25 204- 246 BL00284A 15.64 7.750e-22 73-97 BL00284E 19.15 8.826e-19 391-416 BL00284D 16.34 5.655e- 17 310-337 BL00284B 17.99 3.667e-14 177-198 BL00284A 15.64 2.600e-11 107-131
1170	BL00269	Mammalian defensins proteins.	BL00269C 16.52 6.786e-26 103-

Table 3A  
326

SEQ ID	Database entry ID	Description	Results*
			132 BL00269A 8.53 2.607e-20 38-58 BL00269B 19.17 5.500e-17 65-94
1172	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383A 13.34 5.235e-16 274-289 BL00383F 15.51 1.429e-15 477-493 BL00383E 10.35 7.000e-15 439-450 BL00383D 11.92 4.000e-13 403-416 BL00383C 10.10 6.344e-10 326-337
1172	PR00700	PROTEIN TYROSINE PHOSPHATASE SIGNATURE	PR00700B 16.80 5.320e-21 313-334 PR00700D 12.47 9.217e-19 436-455 PR00700C 13.17 3.143e-15 399-417 PR00700F 11.18 8.941e-13 483-494 PR00700A 6.96 8.714e-10 295-303 PR00700E 17.57 7.618e-09 467-483
1174	BL00472	Small cytokines (intercrine/chemokine) C-C subfamily signatur.	BL00472A 7.45 3.483e-09 1-13
1175	BL00472	Small cytokines (intercrine/chemokine) C-C subfamily signatur.	BL00472A 7.45 3.483e-09 1-13
1187	PR00360	C2 DOMAIN SIGNATURE	PR00360B 13.61 4.273e-09 88-102
1188	PR00860	VERTEBRATE METALLOTHIONEIN SIGNATURE	PR00860B 7.04 6.776e-09 52-66
1189	BL00280	Pancreatic trypsin inhibitor (Kunitz) family proteins.	BL00280 24.61 1.514e-30 276-320 BL00280 24.61 6.553e-24 238-282
1189	PR00759	BASIC PROTEASE (KUNITZ-TYPE) INHIBITOR FAMILY SIGNATURE	PR00759C 14.15 2.385e-14 304-320 PR00759B 11.26 8.105e-12 256-267 PR00759A 11.26 2.452e-10 294-305
1189	BL00213	Lipocalin proteins.	BL00213A 12.95 2.636e-09 38-52
1189	PR00179	LIPOCALIN SIGNATURE	PR00179B 9.56 8.125e-12 122-135 PR00179A 13.78 6.760e-10 38-51 PR00179C 19.02 6.727e-09 150-166
1190	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 1.000e-10 83-117
1191	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 1.000e-10 83-117
1198	BL00154	E1-E2 ATPases phosphorylation site proteins.	BL00154G 21.18 9.783e-09 203-237
1200	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 3.494e-09 120-133
1201	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 8.500e-28 35-74
1202	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 8.500e-28 35-74
1204	BL00134	Serine proteases, trypsin family, histidine proteins.	BL00134A 11.96 7.158e-19 81-98
1204	PR00722	CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1) SIGNATURE	PR00722A 12.27 7.353e-18 82-98
1204	BL00495	Apple domain proteins.	BL00495K 12.58 4.543e-11 83-116
1204	BL01253	Type I fibronectin domain proteins.	BL01253E 16.01 3.415e-13 154-191 BL01253D 4.84 9.585e-11 81-95
1204	BL00021	Kringle domain proteins.	BL00021B 13.33 2.370e-17 81-99

Table 3A  
327

SEQ ID	Database entry ID	Description	Results*
			BL00021C 22.21 6.604e-09 165-187
1208	PR00179	LIPOCALIN SIGNATURE	PR00179B 9.56 7.214e-09 121-134
1212	PR00364	DISEASE RESISTANCE PROTEIN SIGNATURE	PR00364D 10.54 6.700e-09 172-189
1212	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 8.615e-13 177-191 PR00019B 11.36 2.286e-12 174-188 PR00019A 11.19 4.913e-10 319-333 PR00019B 11.36 4.960e-09 387-401 PR00019A 11.19 8.667e-09 151-165
1213	BL00132	Zinc carboxypeptidases, zinc-binding region 1 proteins.	BL00132C 21.35 3.308e-28 129-170 BL00132B 15.93 1.871e-16 99-113 BL00132A 26.07 1.682e-14 50-91 BL00132F 13.26 7.254e-14 228-250 BL00132D 12.70 2.875e-12 173-188 BL00132E 17.72 3.552e-12 199-226
1213	PR00765	CARBOXYPEPTIDASE A METALLOPROTEASE (M14) FAMILY SIGNATURE	PR00765B 15.57 7.857e-16 99-114 PR00765D 14.16 5.500e-11 233-247 PR00765C 12.55 1.290e-10 179-188
1214	PD01270	RECEPTOR FC IMMUNOGLOBULIN AFFIN.	PD01270C 19.54 5.329e-09 32-61
1221	PR00007	COMPLEMENT C1Q DOMAIN SIGNATURE	PR00007A 19.33 1.000e-16 143-170 PR00007C 15.60 8.200e-15 214-236 PR00007B 14.16 5.846e-14 170-190 PR00007D 9.64 5.250e-10 249-260
1221	BL01113	C1q domain proteins.	BL01113B 18.26 1.581e-29 149-185 BL01113C 13.18 3.077e-15 214-234 BL01113A 17.99 1.243e-13 50-77 BL01113A 17.99 6.108e-13 35-62 BL01113A 17.99 3.077e-12 41-68 BL01113A 17.99 1.574e-10 38-65 BL01113A 17.99 9.617e-10 44-71 BL01113A 17.99 7.577e-09 59-86
1221	BL00420	Speract receptor repeat proteins domain proteins.	BL00420A 20.42 2.038e-12 86-115 BL00420A 20.42 5.154e-12 44-73 BL00420A 20.42 4.185e-09 47-76 BL00420A 20.42 9.031e-09 50-79
1223	BL00240	Receptor tyrosine kinase class III proteins.	BL00240E 11.56 2.043e-21 260-298 BL00240F 17.74 4.035e-18 313-361 BL00240G 28.45 4.923e-17 361-414
1223	BL50001	Src homology 2 (SH2) domain proteins profile.	BL50001B 17.40 3.077e-16 271-292
1223	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 4.500e-19 274-305 BL00107B 13.31 7.231e-12 342-358
1223	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109D 17.04 4.462e-20 343-366 PR00109E 14.41 5.950e-20 388-411 PR00109B 12.27 9.571e-19 274-293 PR00109A 15.00 5.065e-10 237-251 PR00109C

Table 3A  
328

SEQ ID	Database entry ID	Description	Results*
			12.85 5.333e-10 324-335
1223	BL00239	Receptor tyrosine kinase class II proteins.	BL00239E 17.14 1.281e-25 314-364 BL00239F 28.15 9.684e-19 369-414 BL00239C 18.75 8.765e-15 261-284 BL00239D 16.81 5.571e-10 286-312
1223	PR00046	MAJOR SIGMA-70 FACTOR SIGNATURE	PR00046A 13.26 8.788e-09 296-310
1223	BL00790	Receptor tyrosine kinase class V proteins.	BL00790O 7.68 9.357e-16 321-354 BL00790Q 15.61 6.057e-11 380-429 BL00790N 13.25 3.937e-10 287-314 BL00790M 8.74 5.138e-10 265-287 BL00790K 9.30 1.000e-08 182-236
1224	PR00453	VON WILLEBRAND FACTOR TYPE A DOMAIN SIGNATURE	PR00453A 12.79 9.827e-09 50-68
1227	BL00615	C-type lectin domain proteins.	BL00615B 12.25 9.571e-10 151-165 BL00615A 16.68 3.100e-09 48-66
1237	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514E 14.28 7.750e-12 313-330
1257	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 1.000e-09 247-261 PR00019B 11.36 2.080e-09 316-330 PR00019B 11.36 3.160e-09 244-258 PR00019B 11.36 4.240e-09 100-114 PR00019B 11.36 5.680e-09 148-162
1263	BL00427	Disintegrins proteins.	BL00427 13.93 4.035e-17 436-491
1263	PR00289	DISINTEGRIN SIGNATURE	PR00289A 13.62 5.667e-14 450-470
1263	PR00873	ECHINOIDEA (SEA URCHIN) METALLOTHIONEIN SIGNATURE	PR00873D 8.43 5.989e-09 608-627
1264	BL01173	Lipolytic enzymes G-D-X-G family, histidine.	BL01173C 8.98 1.000e-16 216-230 BL01173B 13.27 1.000e-12 174-201 BL01173A 9.41 7.200e-11 141-154
1266	BL00492	Clusterin proteins.	BL00492F 16.61 1.000e-40 163-205 BL00492G 13.20 1.000e-40 230-280 BL00492H 16.50 6.870e-40 358-395 BL00492C 8.35 1.000e-38 52-86 BL00492B 10.36 9.053e-28 26-49 BL00492E 12.52 6.423e-26 73-122 BL00492A 11.53 4.724e-18 2-19
1267	BL00284	Serpins proteins.	BL00284C 28.56 6.538e-29 207-249 BL00284A 15.64 3.739e-18 107-131 BL00284D 16.34 3.793e-17 314-341 BL00284E 19.15 2.909e-15 401-426
1268	BL00284	Serpins proteins.	BL00284C 28.56 6.538e-29 243-285 BL00284A 15.64 3.739e-18 125-149 BL00284E 19.15 8.269e-18 437-462 BL00284D 16.34 3.793e-17 350-377 BL00284B 17.99 3.483e-13 214-235
1273	BL00380	Rhodanese proteins.	BL00380D 15.90 8.200e-28 110-



Table 3A  
329

SEQ ID	Database entry ID	Description	Results*
			136 BL00380G 11.26 5.800e-16 251-264 BL00380B 14.77 7.000e-14 49-62 BL00380C 15.67 7.387e-13 82-98 BL00380E 12.44 7.000e-11 181-193 BL00380A 10.48 1.000e-09 10-20
1274	BL00134	Serine proteases, trypsin family, histidine proteins.	BL00134C 13.45 9.000e-17 333-347 BL00134A 11.96 7.429e-16 147-164 BL00134B 15.99 3.793e-15 297-321
1274	BL01253	Type I fibronectin domain proteins.	BL01253H 13.15 5.909e-19 315-350 BL01253G 11.34 8.791e-13 296-310 BL01253F 14.35 2.241e-12 252-291 BL01253E 16.01 7.194e-10 215-252 BL01253D 4.84 7.686e-10 147-161
1274	BL00495	Apple domain proteins.	BL00495N 11.04 3.957e-24 289-324 BL00495O 13.75 6.610e-16 324-353 BL00495K 12.58 1.957e-11 149-182 BL00495L 11.94 2.154e-09 178-217
1274	PR00722	CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1) SIGNATURE	PR00722A 12.27 7.480e-15 148-164 PR00722C 10.87 2.286e-14 296-309 PR00722B 12.51 3.250e-09 203-218
1274	BL00021	Kringle domain proteins.	BL00021D 24.56 8.773e-27 305-347 BL00021B 13.33 9.217e-17 147-165 BL00021C 22.21 3.717e-09 226-248
1279	PR00704	CALPAIN CYSTEINE PROTEASE (C2) FAMILY SIGNATURE	PR00704I 9.52 4.682e-14 145-174
1287	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 8.313e-28 16-55
1287	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 3.769e-15 191-204 PD00066 13.92 1.500e-13 247-260 PD00066 13.92 1.500e-13 387-400 PD00066 13.92 7.500e-13 275-288 PD00066 13.92 8.000e-13 359-372 PD00066 13.92 9.143e-12 219-232 PD00066 13.92 4.913e-11 303-316
1287	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 4.857e-16 371-388 BL00028 16.07 8.412e-14 203-220 BL00028 16.07 5.500e-13 287-304 BL00028 16.07 1.346e-11 259-276 BL00028 16.07 5.154e-11 399-416 BL00028 16.07 7.577e-11 343-360 BL00028 16.07 1.000e-10 175-192 BL00028 16.07 6.700e-10 231-248 BL00028 16.07 3.829e-09 315-332
1287	PD02462	PROTEIN BOLA TRANSCRIPTION REGULATION AC.	PD02462A 22.48 6.817e-09 280-315
1287	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 7.750e-16 368-382 PR00048A 10.52 4.273e-15 284-298 PR00048A 10.52 1.643e-13 200-214 PR00048B 6.02 1.900e-13 188-198 PR00048B 6.02

Table 3A  
330

SEQ ID	Database entry ID	Description	Results*
			6.400e-13 272-282 PR00048A 10.52 3.118e-12 396-410 PR00048B 6.02 5.000e-12 384-394 PR00048B 6.02 4.462e-11 244-254 PR00048B 6.02 1.563e-10 356-366 PR00048A 10.52 2.174e-10 256-270 PR00048A 10.52 4.913e-10 312-326 PR00048A 10.52 9.217e-10 340-354 PR00048B 6.02 3.368e-09 300-310 PR00048A 10.52 7.480e-09 228-242
1292	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243H 17.53 8.696e-11 42-68
1295	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 3.512e-31 10-49
1295	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 9.100e-16 337-350 PD00066 13.92 4.462e-15 477-490 PD00066 13.92 5.800e-14 421-434 PD00066 13.92 9.400e-14 449-462 PD00066 13.92 9.500e-13 365-378 PD00066 13.92 9.500e-13 393-406
1295	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.353e-14 321-338 BL00028 16.07 7.750e-13 405-422 BL00028 16.07 9.550e-13 349-366 BL00028 16.07 9.550e-13 377-394 BL00028 16.07 2.957e-12 433-450 BL00028 16.07 5.800e-10 461-478 BL00028 16.07 7.600e-10 489-506
1295	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE.	PR00048A 10.52 1.000e-13 346-360 PR00048A 10.52 1.000e-13 486-500 PR00048A 10.52 5.500e-13 430-444 PR00048A 10.52 6.786e-13 402-416 PR00048A 10.52 6.786e-13 458-472 PR00048A 10.52 4.789e-11 318-332 PR00048A 10.52 9.526e-11 374-388 PR00048B 6.02 1.000e-10 334-344 PR00048B 6.02 9.438e-10 418-428 PR00048B 6.02 9.438e-10 474-484 PR00048B 6.02 8.579e-09 446-456
1296	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 8.615e-33 16-55
1296	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 6.400e-16 348-361 PD00066 13.92 3.077e-15 236-249 PD00066 13.92 4.462e-15 264-277 PD00066 13.92 4.462e-15 320-333 PD00066 13.92 2.800e-14 432-445 PD00066 13.92 4.600e-14 292-305 PD00066 13.92 8.200e-14 488-501 PD00066 13.92 8.200e-14 516-529 PD00066 13.92 6.500e-13 376-389 PD00066 13.92 7.000e-13 208-221 PD00066 13.92 4.429e-12 460-473
1296	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 9.182e-15 385-399 PR00048A 10.52 9.182e-15

Table 3A  
331

SEQ ID	Database entry ID	Description	Results*
			469-483 PR00048A 10.52 4.750e-14 329-343 PR00048A 10.52 3.571e-13 413-427 PR00048A 10.52 9.357e-13 245-259 PR00048A 10.52 1.529e-12 525-539 PR00048A 10.52 2.588e-12 189-203 PR00048A 10.52 3.118e-12 301-315 PR00048A 10.52 3.647e-12 357-371 PR00048B 6.02 1.000e-11 205-215 PR00048B 6.02 1.000e-11 261-271 PR00048B 6.02 1.000e-11 429-439 PR00048B 6.02 2.385e-11 345-355 PR00048B 6.02 4.462e-11 513-523 PR00048B 6.02 7.231e-11 317-327 PR00048A 10.52 9.526e-11 217-231 PR00048A 10.52 2.174e-10 273-287 PR00048B 6.02 2.688e-10 485-495 PR00048B 6.02 3.813e-10 289-299 PR00048B 6.02 4.938e-10 233-243 PR00048A 10.52 6.478e-10 497-511 PR00048B 6.02 3.842e-09 373-383
1296	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.588e-14 192-209 BL00028 16.07 1.450e-13 388-405 BL00028 16.07 1.450e-13 528-545 BL00028 16.07 5.304e-12 472-489 BL00028 16.07 6.478e-12 416-433 BL00028 16.07 8.826e-12 332-349 BL00028 16.07 9.609e-12 248-265 BL00028 16.07 2.385e-11 276-293 BL00028 16.07 3.769e-11 360-377 BL00028 16.07 7.577e-11 304-321 BL00028 16.07 6.100e-10 220-237 BL00028 16.07 5.886e-09 500-517
1299	BL00142	Neutral zinc metallopeptidases, zinc-binding region proteins.	BL00142 8.38 4.375e-10 188-199
1299	PR00138	MATRIXIN SIGNATURE	PR00138D 16.56 1.321e-19 188-214 PR00138B 15.82 1.655e-10 106-122 PR00138C 16.41 2.393e-10 132-161 PR00138E 6.01 6.586e-10 222-236
1299	BL00024	Hemopexin domain proteins.	BL00024D 17.28 3.411e-22 182-214 BL00024B 21.53 8.788e-19 80-114 BL00024H 11.35 3.077e-10 337-349 BL00024E 7.58 9.211e-10 222-236 BL00024C 22.98 6.000e-09 131-180
1299	PR00480	ASTACIN FAMILY SIGNATURE	PR00480B 15.41 6.529e-09 183-202
1299	BL00546	Matrixins cysteine switch.	BL00546C 16.41 1.643e-22 182-214 BL00546D 10.34 8.017e-10 222-236 BL00546B 20.11 2.443e-09 132-176 BL00546H 10.76 4.512e-09 337-348 BL00546G 16.84 6.684e-09 338-358

Table 3A  
332

SEQ ID	Database entry ID	Description	Results*
1301	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290B 13.17 2.875e-09 170-188
1305	PD02811	PROTEIN PEPTIDE REDUCTASE MG448 PILB FIMBRIA TRAN.	PD02811A 20.67 3.017e-22 60-98 PD02811B 17.07 2.263e-21 111-144 PD02811C 13.25 5.696e-13 147-160
1307	BL00282	Kazal serine protease inhibitors family proteins.	BL00282 16.88 8.043e-18 28-51
1307	PR00290	KAZAL-TYPE SERINE PROTEASE INHIBITOR SIGNATURE	PR00290B 9.78 6.559e-11 39-51 PR00290A 10.88 8.851e-09 28-39
1308	BL00112	ATP:guanido phosphotransferases proteins.	BL00112F 21.37 7.728e-09 51-102
1309	BL00269	Mammalian defensins proteins.	BL00269C 16.52 6.786e-26 139-168 BL00269A 8.53 2.607e-20 74-94 BL00269B 19.17 5.500e-17 101-130
1312	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 3.559e-12 63-107
1312	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 7.545e-17 63-85 PR00449E 13.50 9.280e-14 201-224 PR00449D 10.79 7.158e-13 167-181 PR00449C 17.27 9.229e-12 103-126 PR00449B 14.34 9.280e-10 86-103
1314	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.059e-14 361-378 BL00028 16.07 7.353e-14 193-210 BL00028 16.07 1.900e-13 165-182 BL00028 16.07 3.250e-13 333-350 BL00028 16.07 4.150e-13 221-238 BL00028 16.07 9.550e-13 277-294 BL00028 16.07 1.000e-11 305-322 BL00028 16.07 1.600e-10 249-266
1314	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 2.800e-14 321-334 PD00066 13.92 1.000e-13 265-278 PD00066 13.92 1.429e-12 237-250 PD00066 13.92 7.857e-12 181-194 PD00066 13.92 2.174e-11 349-362 PD00066 13.92 2.385e-10 209-222 PD00066 13.92 7.231e-10 293-306
1314	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 5.500e-14 218-232 PR00048A 10.52 5.500e-14 358-372 PR00048A 10.52 8.500e-14 330-344 PR00048A 10.52 2.286e-13 190-204 PR00048A 10.52 8.071e-13 162-176 PR00048A 10.52 3.118e-12 274-288 PR00048A 10.52 6.684e-11 246-260 PR00048B 6.02 7.231e-11 318-328 PR00048B 6.02 1.563e-10 178-188 PR00048A 10.52 7.261e-10 302-316 PR00048B 6.02 8.579e-09 262-272
1315	PD02283	PROTEIN SPORULATION REPEAT PRECU.	PD02283C 17.54 1.947e-10 95-123
1315	PR00874	FUNGI-IV METALLOTHIONEIN SIGNATURE	PR00874C 4.37 3.750e-09 48-63

Table 3A  
333

SEQ ID	Database entry ID	Description	Results*
1315	BL00269	Mammalian defensins proteins.	BL00269C 16.52 4.434e-09 50-79
1315	PD00866	GLYCOPROTEIN PROTEIN SPIKE E2 PRECURSOR PEPLIMER.	PD00866L 3.73 2.770e-09 24-34 PD00866L 3.73 2.770e-09 82-92 PD00866L 3.73 2.918e-09 4-14 PD00866L 3.73 4.984e-09 14-24
1315	PR00876	NEMATODE METALLOTHIONEIN SIGNATURE	PR00876B 7.66 5.014e-09 89-103
1315	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 5.655e-13 3-46 BL00243I 31.77 9.286e-11 81-124 BL00243I 31.77 6.691e-10 13-56 BL00243I 31.77 2.775e-09 10-53 BL00243I 31.77 5.310e-09 52-95 BL00243I 31.77 5.437e-09 18-61
1315	PR00858	CRUSTACEAN METALLOTHIONEIN SIGNATURE	PR00858B 5.93 3.394e-09 49-68 PR00858B 5.93 7.223e-09 15-34
1315	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 9.893e-13 72-118 BL00203 13.94 7.414e-12 34-80 BL00203 13.94 2.879e-11 46-92 BL00203 13.94 4.956e-11 87-133 BL00203 13.94 5.945e-11 101-147 BL00203 13.94 8.319e-11 78-124 BL00203 13.94 1.383e-10 63-109 BL00203 13.94 3.106e-10 77-123 BL00203 13.94 6.266e-10 15-61 BL00203 13.94 8.085e-10 33-79 BL00203 13.94 8.372e-10 102-148 BL00203 13.94 9.138e-10 56-102 BL00203 13.94 9.521e-10 88-134 BL00203 13.94 1.000e-09 5-51 BL00203 13.94 1.827e-09 82-128 BL00203 13.94 3.847e-09 83-129 BL00203 13.94 5.133e-09 48-94 BL00203 13.94 5.500e-09 25-71 BL00203 13.94 6.143e-09 28-74 BL00203 13.94 6.327e-09 93-139 BL00203 13.94 7.337e-09 73-119 BL00203 13.94 8.071e-09 61-107 BL00203 13.94 8.439e-09 91-137 BL00203 13.94 9.082e-09 13-59
1315	DM00864	EGF-LIKE DOMAIN.	DM00864A 15.21 9.229e-09 6-27
1315	BL00198	4Fe-4S ferredoxins, iron-sulfur binding region proteins.	BL00198 10.43 1.600e-09 23-35 BL00198 10.43 9.700e-09 3-15
1316	DM00406	GLIADIN.	DM00406 7.73 9.514e-10 162-175
1316	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 8.859e-09 66-109
1316	BL00269	Mammalian defensins proteins.	BL00269C 16.52 4.553e-09 24-53 BL00269C 16.52 9.289e-09 94-123
1316	PD02283	PROTEIN SPORULATION REPEAT PRECU.	PD02283C 17.54 3.700e-09 45-73 PD02283C 17.54 8.763e-09 65-93 PD02283C 17.54 9.325e-09 80-108 PD02283C 17.54 9.550e-09 90-118
1316	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 4.363e-11 47-93 BL00203 13.94 5.846e-11 88-134 BL00203 13.94 7.527e-11 57-103 BL00203 13.94 8.714e-11 22-68 BL00203 13.94 4.447e-10 87-133

Table 3A  
334

SEQ ID	Database entry ID	Description	Results*
			BL00203 13.94 7.798e-10 92-138 BL00203 13.94 2.010e-09 37-83 BL00203 13.94 3.939e-09 38-84 BL00203 13.94 5.133e-09 18-64 BL00203 13.94 5.224e-09 17-63 BL00203 13.94 5.592e-09 52-98 BL00203 13.94 5.684e-09 68-114 BL00203 13.94 7.153e-09 33-79 BL00203 13.94 7.245e-09 78-124 BL00203 13.94 8.163e-09 43-89 BL00203 13.94 8.439e-09 23-69 BL00203 13.94 9.265e-09 42-88 BL00203 13.94 9.633e-09 93-139
1317	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 8.279e-10 216-259 BL00243I 31.77 5.310e-09 219-262
1317	BL01208	VWFC domain proteins.	BL01208B 15.83 5.865e-09 244-259
1317	PD02283	PROTEIN SPORULATION REPEAT PRECU.	PD02283C 17.54 7.750e-10 163-191 PD02283C 17.54 4.938e-09 198-226 PD02283C 17.54 6.175e-09 243-271 PD02283C 17.54 6.400e-09 62-90 PD02283C 17.54 6.625e-09 153-181 PD02283C 17.54 6.738e-09 173-201 PD02283C 17.54 6.738e-09 208-236 PD02283C 17.54 7.975e-09 188-216 PD02283C 17.54 9.325e-09 178-206
1317	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 9.703e-11 130-176 BL00203 13.94 1.191e-10 104-150 BL00203 13.94 2.149e-10 64-110 BL00203 13.94 4.447e-10 200-246 BL00203 13.94 5.883e-10 215-261 BL00203 13.94 5.979e-10 210-256 BL00203 13.94 5.979e-10 225-271 BL00203 13.94 6.362e-10 120-166 BL00203 13.94 6.649e-10 24-70 BL00203 13.94 7.032e-10 34-80 BL00203 13.94 8.181e-10 165-211 BL00203 13.94 8.947e-10 85-131 BL00203 13.94 9.809e-10 60-106 BL00203 13.94 1.000e-09 170-216 BL00203 13.94 1.367e-09 176-222 BL00203 13.94 1.643e-09 70-116 BL00203 13.94 2.378e-09 239-285 BL00203 13.94 2.469e-09 125-171 BL00203 13.94 3.755e-09 175-221 BL00203 13.94 3.847e-09 19-65 BL00203 13.94 3.847e-09 190-236 BL00203 13.94 4.122e-09 201-247 BL00203 13.94 4.490e-09 160-206 BL00203 13.94 5.224e-09 186-232 BL00203 13.94 5.867e-09 115-161 BL00203 13.94 6.235e-09 80-126 BL00203 13.94 6.418e-09 55-101 BL00203 13.94 7.337e-09 15-61

Table 3A  
335

SEQ ID	Database entry ID	Description	Results*
			BL00203 13.94 7.337e-09 69-115 BL00203 13.94 7.520e-09 45-91 BL00203 13.94 7.888e-09 95-141 BL00203 13.94 8.255e-09 29-75 BL00203 13.94 8.439e-09 206-252 BL00203 13.94 8.531e-09 126-172 BL00203 13.94 8.806e-09 38-84 BL00203 13.94 8.898e-09 151-197 BL00203 13.94 9.265e-09 216-262 BL00203 13.94 9.449e-09 180-226
1317	PR00876	NEMATODE METALLOTHIONEIN SIGNATURE	PR00876B 7.66 9.757e-09 113-127
1317	BL00269	Mammalian defensins proteins.	BL00269C 16.52 6.425e-10 132-161 BL00269C 16.52 9.882e-09 21-50
1318	PR00876	NEMATODE METALLOTHIONEIN SIGNATURE	PR00876B 7.66 2.268e-10 107-121
1318	PR00874	FUNGI-IV METALLOTHIONEIN SIGNATURE	PR00874C 4.37 9.739e-10 95-110
1318	BL00427	Disintegrins proteins.	BL00427 13.93 2.274e-09 114-169
1318	PD00866	GLYCOPROTEIN PROTEIN SPIKE E2 PRECURSOR PEPLIMER.	PD00866L 3.73 6.564e-10 1-11 PD00866L 3.73 1.443e-09 33-43 PD00866L 3.73 2.770e-09 119-129 PD00866L 3.73 2.918e-09 12-22 PD00866L 3.73 2.918e-09 19-29
1318	BL01208	VWFC domain proteins.	BL01208B 15.83 2.946e-09 143-158
1318	DM01724	kw ALLERGEN POLLEN CIM1 HOL-LI.	DM01724 8.14 7.465e-10 17-37 DM01724 8.14 4.434e-09 19-39 DM01724 8.14 6.684e-09 10-30
1318	PD02283	PROTEIN SPORULATION REPEAT PRECU.	PD02283C 17.54 8.763e-09 132-160
1318	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 1.724e-12 124-170 BL00203 13.94 2.034e-12 119-165 BL00203 13.94 4.207e-12 115-161 BL00203 13.94 4.931e-12 93-139 BL00203 13.94 3.275e-11 109-155 BL00203 13.94 3.967e-11 120-166 BL00203 13.94 3.298e-10 114-160 BL00203 13.94 4.064e-10 35-81 BL00203 13.94 4.447e-10 100-146 BL00203 13.94 5.117e-10 98-144 BL00203 13.94 7.989e-10 83-129 BL00203 13.94 9.426e-10 125-171 BL00203 13.94 2.929e-09 34-80 BL00203 13.94 2.929e-09 94-140 BL00203 13.94 3.663e-09 32-78 BL00203 13.94 4.490e-09 88-134 BL00203 13.94 4.582e-09 110-156 BL00203 13.94 4.673e-09 101-147 BL00203 13.94 4.857e-09 104-150 BL00203 13.94 5.592e-09 43-89 BL00203 13.94 6.694e-09 111-157 BL00203 13.94 8.898e-09 55-101 BL00203 13.94 9.082e-09 99-145
1318	BL00243	Integrins beta chain cysteine-rich	BL00243I 31.77 1.662e-10 68-111

Table 3A  
336

SEQ ID	Database entry ID	Description	Results*
		domain proteins.	BL00243I 31.77 3.912e-10 38-81 BL00243I 31.77 5.500e-10 58-101 BL00243I 31.77 8.941e-10 48-91 BL00243I 31.77 6.324e-09 25-68 BL00243I 31.77 7.465e-09 118-161 BL00243I 31.77 9.239e-09 61-104 BL00243I 31.77 9.620e-09 98-141
1319	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 2.800e-14 107-120 PD00066 13.92 2.800e-14 219-232 PD00066 13.92 4.000e-13 79-92 PD00066 13.92 4.000e-12 51-64 PD00066 13.92 9.143e-12 191-204 PD00066 13.92 2.038e-10 163-176
1319	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 3.250e-13 203-220 BL00028 16.07 8.650e-13 35-52 BL00028 16.07 9.550e-13 91-108 BL00028 16.07 3.739e-12 147-164 BL00028 16.07 4.522e-12 119-136 BL00028 16.07 4.462e-11 175-192 BL00028 16.07 8.500e-10 63-80
1319	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 4.706e-12 60-74 PR00048B 6.02 6.000e-12 104-114 PR00048B 6.02 1.000e-11 216-226 PR00048A 10.52 3.368e-11 88-102 PR00048B 6.02 7.231e-11 160-170 PR00048A 10.52 7.632e-11 32-46 PR00048B 6.02 1.000e-10 188-198 PR00048A 10.52 2.565e-10 172-186 PR00048A 10.52 2.957e-10 200-214 PR00048B 6.02 1.000e-09 48-58 PR00048A 10.52 4.240e-09 144-158 PR00048B 6.02 6.684e-09 76-86 PR00048B 6.02 7.632e-09 132-142 PR00048A 10.52 8.560e-09 116-130
1319	PD02462	PROTEIN BOLA TRANSCRIPTION REGULATION AC.	PD02462A 22.48 9.122e-09 168-203
1320	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 8.800e-15 4-26
1320	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 3.250e-09 4-48
1321	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 5.055e-11 71-117 BL00203 13.94 8.439e-09 146-192 BL00203 13.94 9.082e-09 53-99
1327	BL00799	Granulins proteins.	BL00799H 14.15 9.667e-09 18-59
1329	PF00711	Beta defensins.	PF00711 15.76 7.915e-11 60-93
1329	PR00858	CRUSTACEAN METALLOTHIONEIN SIGNATURE	PR00858B 5.93 1.479e-09 55-74
1329	BL00317	WAP-type 'four-disulfide core' domain proteins.	BL00317B 14.58 2.216e-09 63-85
1329	DM01724	kw ALLERGEN POLLEN CIM1 HOLLI.	DM01724 8.14 4.484e-12 19-39 DM01724 8.14 4.484e-12 23-43 DM01724 8.14 4.484e-12 27-47 DM01724 8.14 4.484e-12 31-51 DM01724 8.14 4.296e-10 35-55 DM01724 8.14 4.197e-09 11-31



Table 3A  
337

SEQ ID	Database entry ID	Description	Results*
			DM01724 8.14 5.382e-09 7-27 DM01724 8.14 6.803e-09 15-35
1329	PR00872	DIPTERA (DROSOPHILA) METALLOTHIONEIN SIGNATURE	PR00872A 5.17 7.409e-09 99-112
1329	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 7.429e-13 47-93 BL00203 13.94 1.989e-11 54-100 BL00203 13.94 3.967e-11 50-96 BL00203 13.94 8.372e-10 51-97 BL00203 13.94 8.947e-10 58-104 BL00203 13.94 1.184e-09 75-121 BL00203 13.94 3.204e-09 74-120 BL00203 13.94 3.663e-09 57-103 BL00203 13.94 4.673e-09 65-111 BL00203 13.94 5.592e-09 70-116 BL00203 13.94 6.786e-09 55-101 BL00203 13.94 7.429e-09 40-86
1329	BL00264	Neurohypophysial hormones proteins.	BL00264 8.98 7.632e-09 94-121
1329	PD00866	GLYCOPROTEIN PROTEIN SPIKE E2 PRECURSOR PEPLIMER.	PD00866L 3.73 7.709e-10 74-84 PD00866L 3.73 6.902e-09 4-14 PD00866L 3.73 9.262e-09 1-11
1329	DM00864	EGF-LIKE DOMAIN.	DM00864A 15.21 9.486e-09 9-30
1329	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 2.000e-11 57-100 BL00243I 31.77 5.897e-10 69-112 BL00243I 31.77 1.254e-09 60-103 BL00243I 31.77 7.465e-09 8-51 BL00243I 31.77 1.000e-08 16-59
1330	BL00412	Neuromodulin (GAP-43) proteins.	BL00412D 16.54 2.957e-10 175-226
1330	DM01206	CORONAVIRUS NUCLEOCAPSID PROTEIN.	DM01206B 10.69 3.278e-09 70-90 DM01206B 10.69 4.418e-09 105-125
1330	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 5.291e-09 125-160
1330	PF00992	Troponin.	PF00992A 16.67 5.026e-09 183-218 PF00992A 16.67 8.934e-09 191-226
1330	BL00048	Protamine P1 proteins.	BL00048 6.39 7.107e-16 64-91 BL00048 6.39 9.196e-16 63-90 BL00048 6.39 1.132e-12 62-89 BL00048 6.39 2.059e-12 66-93 BL00048 6.39 3.250e-12 65-92 BL00048 6.39 7.618e-12 92-119 BL00048 6.39 2.625e-11 60-87 BL00048 6.39 6.500e-11 113-140 BL00048 6.39 6.750e-11 78-105 BL00048 6.39 6.875e-11 104-131 BL00048 6.39 7.125e-11 112-139 BL00048 6.39 8.625e-11 74-101 BL00048 6.39 2.539e-10 108-135 BL00048 6.39 4.434e-10 61-88 BL00048 6.39 5.855e-10 110-137 BL00048 6.39 6.921e-10 98-125 BL00048 6.39 7.158e-10 109-136 BL00048 6.39 7.750e-10 97-124 BL00048 6.39 8.105e-10 79-106 BL00048 6.39 8.579e-10 19-46

Table 3A  
338

SEQ ID	Database entry ID	Description	Results*
			BL00048 6.39 8.934e-10 94-121 BL00048 6.39 9.526e-10 103-130 BL00048 6.39 1.675e-09 101-128 BL00048 6.39 1.900e-09 73-100 BL00048 6.39 3.250e-09 81-108 BL00048 6.39 3.475e-09 111-138 BL00048 6.39 3.700e-09 82-109 BL00048 6.39 3.700e-09 96-123 BL00048 6.39 4.263e-09 99-126 BL00048 6.39 5.163e-09 107-134 BL00048 6.39 5.275e-09 67-94 BL00048 6.39 5.275e-09 80-107 BL00048 6.39 5.388e-09 49-76 BL00048 6.39 6.738e-09 116-143 BL00048 6.39 8.650e-09 52-79 BL00048 6.39 8.763e-09 18-45 BL00048 6.39 9.100e-09 21-48 BL00048 6.39 9.550e-09 76-103 BL00048 6.39 9.550e-09 100-127 BL00048 6.39 9.663e-09 102-129 BL00048 6.39 1.000e-08 77-104
1331	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290B 13.17 7.750e-09 74-92
1332	BL00280	Pancreatic trypsin inhibitor (Kunitz) family proteins.	BL00280 24.61 2.154e-28 83-127
1332	PR00759	BASIC PROTEASE (KUNITZ-TYPE) INHIBITOR FAMILY SIGNATURE	PR00759B 11.26 3.323e-10 101-112 PR00759C 14.15 2.125e-09 111-127
1333	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.162e-34 8-47
1333	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 2.800e-14 273-286 PD00066 13.92 2.800e-14 301-314 PD00066 13.92 2.800e-14 329-342 PD00066 13.92 2.800e-14 357-370 PD00066 13.92 5.500e-13 245-258 PD00066 13.92 1.783e-11 217-230 PD00066 13.92 8.043e-11 189-202
1333	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 6.850e-13 257-274 BL00028 16.07 2.957e-12 201-218 BL00028 16.07 4.522e-12 341-358 BL00028 16.07 5.304e-12 285-302 BL00028 16.07 1.346e-11 369-386 BL00028 16.07 2.038e-11 229-246 BL00028 16.07 4.900e-10 313-330
1333	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 2.286e-13 198-212 PR00048A 10.52 2.929e-13 254-268 PR00048A 10.52 2.929e-13 282-296 PR00048A 10.52 5.500e-13 366-380 PR00048B 6.02 6.400e-13 354-364 PR00048A 10.52 8.941e-12 226-240 PR00048B 6.02 1.000e-11 326-336 PR00048B 6.02 5.846e-11 242-252 PR00048A 10.52 6.211e-11 310-324 PR00048A 10.52 8.579e-11 338-352 PR00048B 6.02 9.308e-11

Table 3A  
339

SEQ ID	Database entry ID	Description	Results*
			270-280 PR00048B 6.02 7.750e-10 298-308 PR00048B 6.02 8.313e-10 214-224
1336	PR00806	VINCULIN SIGNATURE	PR00806B 4.28 8.920e-09 333-347 PR00806B 4.28 9.640e-09 332-346
1337	PR00806	VINCULIN SIGNATURE	PR00806B 4.28 8.920e-09 333-347 PR00806B 4.28 9.640e-09 332-346
1340	BL01221	PMP-22 / EMP / MP20 family proteins.	BL01221B 13.29 6.745e-09 54-68
1343	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 8.200e-12 208-231
1348	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 1.947e-10 316-326
1352	DM00973	3 kw RESISTANCE BENOMYL YLL028W CYCLOHEXIMIDE.	DM00973A 21.17 2.306e-13 66-103
1353	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 8.200e-10 190-204 PR00019A 11.19 7.667e-09 193-207
1356	PD02327	GLYCOPROTEIN ANTIGEN PRECURSOR IMMUNOGLO.	PD02327B 19.84 2.091e-09 144-166
1356	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 7.652e-09 291-301
1357	PF01130	CD36 family.	PF01130D 26.65 1.955e-23 337-391 PF01130C 22.91 1.851e-18 255-305 PF01130B 21.40 1.537e-15 200-246
1359	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 7.480e-10 160-183 BL00290B 13.17 2.875e-09 226-244
1366	PR00007	COMPLEMENT C1Q DOMAIN SIGNATURE	PR00007A 19.33 6.936e-13 274-301 PR00007C 15.60 9.250e-13 349-371 PR00007B 14.16 9.372e-13 301-321 PR00007D 9.64 5.500e-11 381-392
1366	BL01113	C1q domain proteins.	BL01113B 18.26 1.675e-24 280-316 BL01113A 17.99 4.194e-15 200-227 BL01113D 7.47 3.250e-13 383-393 BL01113A 17.99 3.919e-13 191-218 BL01113C 13.18 9.294e-12 349-369 BL01113A 17.99 5.500e-11 185-212 BL01113A 17.99 6.727e-11 182-209 BL01113A 17.99 8.773e-11 203-230 BL01113A 17.99 3.681e-10 188-215 BL01113A 17.99 6.936e-10 176-203 BL01113A 17.99 7.319e-10 194-221 BL01113A 17.99 4.635e-09 209-236 BL01113A 17.99 5.500e-09 179-206
1366	BL00420	Speract receptor repeat proteins domain proteins.	BL00420A 20.42 3.250e-12 185-214 BL00420A 20.42 2.623e-10 203-232 BL00420A 20.42 2.385e-09 191-220 BL00420A 20.42 6.400e-09 182-211 BL00420A 20.42 7.646e-09 188-217 BL00420A 20.42 8.200e-09 173-

Table 3A  
340

SEQ ID	Database entry ID	Description	Results*
			202
1366	PR00524	CHOLECYSTOKININ TYPE A RECEPTOR SIGNATURE	PR00524F 5.36 1.766e-09 200-214 PR00524F 5.36 8.564e-09 188-202
1367	PR00794	PANCREATIC RIBONUCLEASE FAMILY SIGNATURE	PR00794D 17.96 8.800e-17 94-117 PR00794A 14.47 6.897e-16 46-66 PR00794B 16.95 5.655e-14 66-86
1367	BL00127	Pancreatic ribonuclease family proteins.	BL00127B 26.57 9.357e-29 46-91 BL00127C 31.49 9.763e-22 82-126 BL00127A 12.84 8.200e-10 31-41
1368	BL00067	3-hydroxyacyl-CoA dehydrogenase proteins.	BL00067A 21.37 9.053e-10 53-84
1368	BL00677	D-amino acid oxidases proteins.	BL00677A 6.96 3.821e-09 55-68
1368	BL00982	Bacterial-type phytoene dehydrogenase proteins.	BL00982A 18.41 4.000e-09 57-89
1369	BL00067	3-hydroxyacyl-CoA dehydrogenase proteins.	BL00067A 21.37 9.053e-10 53-84
1369	BL00677	D-amino acid oxidases proteins.	BL00677A 6.96 3.821e-09 55-68
1369	BL00982	Bacterial-type phytoene dehydrogenase proteins.	BL00982A 18.41 4.000e-09 57-89
1371	PR00762	CHLORIDE CHANNEL SIGNATURE	PR00762C 9.29 5.200e-20 203-223 PR00762A 14.22 8.364e-19 108-126 PR00762D 11.29 9.379e-19 440-461 PR00762B 12.12 7.750e-16 139-159 PR00762E 12.07 2.286e-15 475-492 PR00762F 15.12 6.559e-15 493-513 PR00762G 14.13 1.692e-11 531-546
1371	PD00866	GLYCOPROTEIN PROTEIN SPIKE E2 PRECURSOR PEPLIMER.	PD00866L 3.73 3.951e-09 9-19
1371	PD00120	REPEAT PRECURSOR GLYCOPROTEIN EG.	PD00120A 12.94 4.750e-09 6-19
1372	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 4.349e-18 100-122
1372	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 2.800e-13 131-171
1372	PR00642	EDG1 ORPHAN RECEPTOR SIGNATURE	PR00642D 9.70 5.065e-12 89-104
1372	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237C 15.69 6.400e-09 145-168
1374	BL01022	PTR2 family proton/oligopeptide symporters proteins.	BL01022B 22.19 2.241e-15 74-120 BL01022A 11.58 2.212e-12 44-63 BL01022C 16.62 4.919e-10 160-184
1374	PR00308	TYPE I ANTIFREEZE PROTEIN SIGNATURE	PR00308C 3.83 2.169e-09 20-30
1375	PR00705	PAPAIN CYSTEINE PROTEASE (C1) FAMILY SIGNATURE	PR00705A 10.55 4.000e-21 132-148 PR00705B 10.22 2.385e-10 276-287
1375	BL00139	Eukaryotic thiol (cysteine) proteases cysteine proteins.	BL00139D 9.24 1.818e-18 295-312 BL00139A 10.29 1.000e-14 132-142 BL00139C 9.23 2.800e-10 275-285
1375	PR00704	CALPAIN CYSTEINE PROTEASE (C2) FAMILY SIGNATURE	PR00704C 11.88 6.162e-09 132-149
1376	PR00237	RHODOPSIN-LIKE GPCR	PR00237C 15.69 3.333e-12 111-

Table 3A

341

SEQ ID	Database entry ID	Description	Results*
		SUPERFAMILY SIGNATURE	134 PR00237E 13.03 5.667e-11 193-217 PR00237F 13.57 1.474e-09 397-422 PR00237B 13.50 4.750e-09 66-88
1376	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 9.625e-19 97-137 BL00237C 13.19 8.846e-09 392-419 BL00237B 5.28 9.182e-09 201-213
1378	BL00649	G-protein coupled receptors family 2 proteins.	BL00649C 17.82 5.773e-10 325-351
1378	PR00249	SECRETIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00249C 17.08 4.130e-09 327-351
1380	PD01876	ANTIGEN MELANOMA-ASSOCIATED MULTIGENE FAMILY TUM.	PD01876A 12.01 2.964e-26 1-42 PD01876C 21.73 2.343e-25 99-152 PD01876B 14.18 2.500e-12 62-96
1382	DM01724	kw ALLERGEN POLLEN CIM1 HOL-LI.	DM01724 8.14 3.789e-10 291-311
1382	PR00213	MYELIN P0 PROTEIN SIGNATURE	PR00213E 5.51 8.969e-09 263-288
1382	PD01849	CHORION CLASS PRECURSOR EGG SHELL REPEAT MULTIGEN.	PD01849C 11.22 9.625e-09 265-293
1388	BL00170	Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatur.	BL00170B 20.97 1.000e-40 88-128 BL00170C 18.49 1.409e-37 135-180 BL00170A 17.08 6.850e-22 58-85
1388	PR00153	CYCLOPHILIN PEPTIDYL-PROLYL CIS-TRANS ISOMERASE SIGNATURE	PR00153C 11.01 2.800e-18 136-152 PR00153B 11.57 8.500e-17 93-106 PR00153D 11.99 4.000e-16 151-164 PR00153A 12.98 4.789e-16 64-80 PR00153E 9.10 2.980e-10 164-180
1390	BL00237	G-protein coupled receptors proteins.	BL00237C 13.19 6.571e-16 147-174 BL00237A 27.68 2.976e-11 13-53 BL00237B 5.28 2.421e-10 102-114
1390	PR00529	GONADOTROPHIN RELEASING HORMONE RECEPTOR SIGNATURE	PR00529D 23.59 7.380e-13 2-18 PR00529H 10.07 1.000e-10 176-197 PR00529F 5.98 5.138e-10 75-90
1390	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237F 13.57 7.429e-15 152-177 PR00237E 13.03 3.250e-14 94-118 PR00237D 8.94 7.000e-09 58-80 PR00237G 19.63 8.116e-09 192-219
1390	PD02448	TRANSCRIPTION PROTEIN DNA-BINDIN.	PD02448A 9.37 9.195e-09 34-73
1392	BL00216	Sugar transport proteins.	BL00216B 27.64 3.732e-11 196-246
1393	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237C 15.69 9.333e-12 143-166
1393	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 3.200e-10 129-169
1393	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 2.429e-12 116-138 PR00245B 10.38 7.525e-11 216-231 PR00245E 12.40 2.019e-10 330-345 PR00245C 7.84 7.286e-10 277-293

Table 3A  
342

SEQ ID	Database entry ID	Description	Results*
1397	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014C 15.44 1.783e-09 585-604 PR00014A 8.22 3.045e-09 747-757 PR00014C 15.44 6.087e-09 683-702
1397	BL00790	Receptor tyrosine kinase class V proteins.	BL00790I 20.01 1.750e-12 1016-1047 BL00790I 20.01 6.125e-12 694-725 BL00790I 20.01 6.679e-09 596-627
1397	PR00096	GLUTAMINE AMIDOTRANSFERASE SUPERFAMILY SIGNATURE	PR00096B 9.72 9.827e-09 1063-1075
1401	PD00866	GLYCOPROTEIN PROTEIN SPIKE E2 PRECURSOR PEPLIMER.	PD00866L 3.73 8.820e-09 51-61
1402	PR00250	FUNGAL PHEROMONE MATING FACTOR STE2 GPCR SIGNATURE	PR00250G 9.58 9.746e-09 21-43
1404	PD01652	RECEPTOR CELL NK GLYCOPROTEIN IMMUNOGLOB.	PD01652A 15.35 6.625e-10 24-60 PD01652B 8.50 1.836e-09 14-66 PD01652B 8.50 4.021e-09 111-163
1409	BL00240	Receptor tyrosine kinase class III proteins.	BL00240B 24.70 8.250e-12 268-292
1409	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 3.842e-10 312-322
1409	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870B 18.83 5.500e-10 421-454 PD02870D 15.74 6.266e-09 182-217
1409	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 8.405e-09 496-509
1414	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 3.455e-18 46-71 BL00215B 10.44 8.714e-12 183-196 BL00215A 15.82 7.319e-11 143-168
1414	PR00927	ADENINE NUCLEOTIDE TRANSLOCATOR 1 SIGNATURE	PR00927E 14.93 6.040e-09 74-96
1414	BL00439	Acyltransferases ChoActase / COT / CPT family proteins.	BL00439A 9.40 8.279e-09 318-335
1416	PD01270	RECEPTOR FC IMMUNOGLOBULIN AFFIN.	PD01270D 24.66 9.286e-12 12-48
1417	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 2.714e-10 42-51
1417	BL01282	BIR repeat proteins.	BL01282B 30.49 5.114e-09 23-62
1418	BL00604	Synaptophysin / synaptoporin proteins.	BL00604A 9.13 7.035e-29 41-96 BL00604E 8.32 4.759e-24 201-243 BL00604C 14.66 8.807e-17 134-166 BL00604B 9.95 5.154e-16 104-134 BL00604D 12.28 5.100e-14 166-201
1418	PR00220	SYNAPTOPHYSIN/SYNAPTOPORIN FAMILY SIGNATURE	PR00220A 10.93 1.353e-21 38-61 PR00220E 3.46 4.150e-18 216-235 PR00220D 8.32 4.575e-15 149-173 PR00220C 11.05 5.557e-14 117-142 PR00220B 15.48 2.703e-13 62-88
1421	BL01219	Ammonium transporters proteins.	BL01219D 11.63 2.957e-10 217-241 BL01219F 15.24 8.809e-09 289-314
1421	PR00342	RHESUS BLOOD GROUP PROTEIN	PR00342G 8.18 1.458e-19 220-239

Table 3A  
343

SEQ ID	Database entry ID	Description	Results*
		SIGNATURE	PR00342B 11.09 8.657e-13 61-79 PR00342D 8.46 2.857e-12 128-145 PR00342H 7.61 6.927e-11 250-273 PR00342C 10.10 4.770e-10 90-108 PR00342E 14.49 5.950e-10 151-175 PR00342F 7.02 1.556e-09 185-201 PR00342L 7.61 9.600e-09 352-378
1423	PR00794	PANCREATIC RIBONUCLEASE FAMILY SIGNATURE	PR00794C 17.32 8.071e-14 96-115 PR00794B 16.95 6.455e-13 71-91 PR00794A 14.47 9.538e-13 51-71 PR00794D 17.96 9.368e-10 117-140
1423	BL00127	Pancreatic ribonuclease family proteins.	BL00127B 26.57 2.895e-23 51-96 BL00127C 31.49 3.045e-11 105-149 BL00127A 12.84 2.161e-09 36-46
1426	PR00830	ENDOPEPTIDASE LA (LON) SERINE PROTEASE (S16) SIGNATURE	PR00830A 8.41 1.143e-11 165-185
1426	BL00113	Adenylate kinase proteins.	BL00113A 12.74 4.462e-09 471-488
1426	BL00674	AAA-protein family proteins.	BL00674B 4.46 8.905e-09 467-489
1427	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 9.500e-11 74-88 PR00019A 11.19 5.696e-10 77-91 PR00019B 11.36 6.400e-10 25-39 PR00019B 11.36 5.320e-09 49-63
1427	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014C 15.44 8.043e-09 382-401
1437	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 6.850e-15 889-917 PD01719A 12.89 6.727e-13 643-671 PD01719A 12.89 1.913e-12 832-860 PD01719A 12.89 4.261e-12 701-729
1442	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 1.900e-11 106-120
1443	BL00353	HMG1/2 proteins.	BL00353B 11.47 1.643e-32 72-122
1443	PR00886	HIGH MOBILITY GROUP (HMG1/HMG2) PROTEIN SIGNATURE	PR00886A 10.08 7.218e-18 37-60 PR00886B 9.88 1.750e-12 60-81 PR00886C 11.84 9.016e-12 104-123
1443	BL00649	G-protein coupled receptors family 2 proteins.	BL00649C 17.82 6.226e-11 286-312
1443	PR00249	SECRETIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00249C 17.08 1.783e-09 288-312 PR00249D 13.09 5.143e-09 327-353
1446	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 2.373e-09 781-814
1448	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 8.780e-09 402-435
1448	PF00761	Polyomavirus coat protein.	PF00761A 12.61 8.925e-09 1-26
1449	PR00180	CELLULAR RETINALDEHYDE-BINDING PROTEIN SIGNATURE	PR00180B 16.42 7.146e-09 236-261
1452	PR00926	MITOCHONDRIAL CARRIER PROTEIN SIGNATURE	PR00926F 17.75 4.706e-24 132-155 PR00926E 11.70 7.000e-19 89-108 PR00926D 10.53 9.308e-17 41-60 PR00926F 17.75 9.859e-10 35-58

Table 3A  
344

SEQ ID	Database entry ID	Description	Results*
1452	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 8.235e-10 33-58 BL00215B 10.44 6.400e-09 79-92 BL00215B 10.44 8.200e-09 174-187
1452	PR00927	ADENINE NUCLEOTIDE TRANSLOCATOR 1 SIGNATURE	PR00927E 14.93 7.652e-22 67-89 PR00927G 11.07 3.500e-15 178-194 PR00927F 11.79 6.211e-14 128-145 PR00927D 11.02 4.000e-13 26-40 PR00927B 14.66 2.597e-12 164-186 PR00927E 14.93 7.818e-11 162-184 PR00927A 7.98 9.667e-09 28-41
1456	BL00239	Receptor tyrosine kinase class II proteins.	BL00239B 25.15 1.000e-40 1037-1085 BL00239E 17.14 1.000e-40 1154-1204 BL00239F 28.15 1.750e-31 1208-1253 BL00239C 18.75 1.000e-24 1099-1122 BL00239D 16.81 2.884e-15 1124-1150
1456	BL00107	Protein kinases ATP-binding region proteins.	BL00107B 13.31 1.900e-14 1182-1198 BL00107A 18.39 9.217e-14 1112-1143
1456	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109D 17.04 4.316e-23 1183-1206 PR00109B 12.27 4.857e-19 1112-1131 PR00109E 14.41 5.500e-16 1227-1250 PR00109C 12.85 4.000e-13 1164-1175 PR00109A 15.00 8.962e-11 1066-1080
1456	BL00790	Receptor tyrosine kinase class V proteins.	BL00790O 7.68 4.532e-18 1161-1194 BL00790K 9.30 5.091e-16 1010-1064 BL00790H 13.42 1.466e-15 830-856 BL00790M 8.74 9.203e-12 1103-1125 BL00790Q 15.61 5.200e-11 1219-1268 BL00790P 12.33 8.057e-10 1194-1219
1456	BL00240	Receptor tyrosine kinase class III proteins.	BL00240F 17.74 3.875e-26 1153-1201 BL00240C 22.58 4.414e-16 978-1027 BL00240E 11.56 5.061e-15 1098-1136 BL00240G 28.45 8.909e-10 1200-1253
1456	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014D 12.04 2.350e-09 893-908 PR00014B 14.77 7.000e-09 842-853
1456	BL50001	Src homology 2 (SH2) domain proteins profile.	BL50001B 17.40 9.640e-09 1109-1130
1457	PF01307	Plant virus coat protein.	PF01307A 21.19 3.025e-09 321-351
1462	BL00546	Matrixins cysteine switch.	BL00546C 16.41 1.643e-33 232-264 BL00546B 20.11 8.895e-29 18-62 BL00546A 19.62 7.667e-26 91-121 BL00546B 20.11 7.720e-19 180-224 BL00546D 10.34 1.000e-15 273-287 BL00546E 10.23 4.529e-15 309-330



Table 3A  
345

SEQ ID	Database entry ID	Description	Results*
1462	BL00024	Hemopexin domain proteins.	BL00024D 17.28 7.375e-34 232-264 BL00024C 22.98 5.091e-32 17-66 BL00024B 21.53 3.143e-23 130-164 BL00024C 22.98 8.732e-22 179-228 BL00024E 7.58 5.500e-15 273-287 BL00024F 11.30 1.900e-14 309-330 BL00024A 11.49 9.100e-13 111-122
1462	BL00142	Neutral zinc metalloproteinases, zinc-binding region proteins.	BL00142 8.38 1.000e-12 238-249
1462	PR00138	MATRIXIN SIGNATURE	PR00138D 16.56 6.500e-30 238-264 PR00138C 16.41 5.846e-16 18-47 PR00138A 15.14 7.136e-16 111-125 PR00138E 6.01 7.000e-15 273-287 PR00138B 15.82 3.824e-11 156-172
1462	PR00480	ASTACIN FAMILY SIGNATURE	PR00480B 15.41 6.045e-10 233-252
1475	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354Z 9.06 9.566e-17 216-265 DM01354Z 9.06 2.534e-15 161-210
1478	PR00185	EUKARYOTIC PORIN SIGNATURE	PR00185D 16.54 1.136e-11 111-129
1482	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449E 13.50 4.971e-11 464-487
1484	PR00879	FISH ACETYLCHOLINESTERASE SIGNATURE	PR00879A 6.28 1.000e-08 39-45
1487	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354Q 9.23 1.000e-40 283-338 DM01354T 9.48 4.882e-39 389-432 DM01354R 8.50 8.380e-28 338-368 DM01354P 9.18 2.957e-26 249-283 DM01354S 11.61 1.000e-18 368-389
1489	PD02327	GLYCOPROTEIN ANTIGEN PRECURSOR IMMUNOGLO.	PD02327C 15.47 3.538e-09 60-75
1491	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354B 14.71 6.426e-14 271-299
1493	PR00502	MUTT DOMAIN SIGNATURE	PR00502B 15.92 4.643e-09 348-364
1494	PF00997	Kappa casein.	PF00997D 9.95 8.306e-09 117-152
1495	DM01162	BARLEY YELLOW DWARF VIRUS RNA-DIRECTED RNA POLYMERASE.	DM01162B 11.50 9.743e-09 104-113
1497	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 9.757e-10 79-113
1498	PR00069	ALDO-KETO REDUCTASE SIGNATURE	PR00069D 19.36 1.659e-19 260-290
1498	BL00798	Aldo/keto reductase family proteins.	BL00798F 23.30 5.125e-25 314-363 BL00798E 20.32 7.462e-24 256-294 BL00798A 14.97 3.000e-10 207-222
1499	BL00826	MARCKS family proteins.	BL00826B 12.51 2.116e-09 545-597
1499	BL00970	Nuclear transition protein 2 proteins.	BL00970C 14.80 5.538e-09 567-605
1499	DM00303	6 LEA 11-MER REPEAT REPEAT.	DM00303A 13.20 8.327e-09 574-

Table 3A  
346

SEQ ID	Database entry ID	Description	Results*
			624
1499	PD01457	RIBOSOMAL PROTEIN 40S ZINC-FINGER METAL.	PD01457A 16.51 7.708e-11 561-606 PD01457A 16.51 1.409e-10 565-610 PD01457A 16.51 5.745e-10 559-604 PD01457A 16.51 2.397e-09 558-603 PD01457A 16.51 3.483e-09 567-612 PD01457A 16.51 4.103e-09 560-605 PD01457A 16.51 8.060e-09 563-608 PD01457A 16.51 8.293e-09 557-602 PD01457A 16.51 8.914e-09 562-607
1499	BL00348	p53 tumor antigen proteins.	BL00348F 23.19 5.571e-09 543-586 BL00348F 23.19 6.429e-09 583-626 BL00348F 23.19 9.286e-09 560-603
1501	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354X 13.86 6.230e-15 3-42
1502	BL01210	Caveolins proteins.	BL01210A 17.61 9.438e-14 132-171
1503	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.571e-10 65-80
1503	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 1.161e-10 53-86 DM00215 19.43 8.393e-10 52-85 DM00215 19.43 4.661e-09 47-80 DM00215 19.43 5.729e-09 46-79
1506	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354Z 9.06 8.548e-13 66-115
1508	PD01941	TRANSMEMBRANE COTRANSORTER SYMP.	PD01941E 15.92 8.714e-23 198-245 PD01941F 28.52 5.154e-11 496-551
1510	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354N 13.17 1.000e-40 148-193 DM01354V 12.97 1.000e-40 274-321 DM01354V 12.97 6.932e-40 218-265 DM01354O 8.73 1.882e-19 199-245 DM01354W 12.64 2.427e-13 265-285 DM01354W 12.64 1.300e-11 321-341
1511	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.913e-19 230-269
1511	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 4.600e-13 343-360 BL00028 16.07 6.087e-12 371-388
1511	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 3.250e-16 340-354 PR00048A 10.52 3.348e-10 368-382 PR00048B 6.02 4.938e-10 384-394
1511	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 3.077e-10 331-344 PD00066 13.92 6.700e-09 359-372
1517	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 8.875e-15 449-466 BL00028 16.07 6.824e-14 309-326 BL00028 16.07 7.353e-14 505-522 BL00028 16.07 2.800e-13 1038-1055 BL00028 16.07 8.650e-13 337-354 BL00028 16.07 1.783e-12 926-943 BL00028 16.07 4.913e-12

Table 3A  
347

SEQ ID	Database entry ID	Description	Results*
			982-999 BL00028 16.07 5.304e-12 898-915 BL00028 16.07 5.304e-12 1094-1111 BL00028 16.07 8.435e-12 421-438 BL00028 16.07 3.077e-11 393-410 BL00028 16.07 3.423e-11 954-971 BL00028 16.07 3.769e-11 365-382 BL00028 16.07 4.462e-11 870-887 BL00028 16.07 5.154e-11 281-298 BL00028 16.07 4.000e-10 477-494 BL00028 16.07 7.000e-10 1010-1027 BL00028 16.07 7.000e-10 1066-1083
1517	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 9.400e-30 76-115 PD01066 19.43 7.525e-15 675-714 PD01066 19.43 1.943e-09 716-755
1517	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 5.500e-16 446-460 PR00048A 10.52 9.182e-15 306-320 PR00048A 10.52 9.182e-15 1035-1049 PR00048A 10.52 5.500e-14 502-516 PR00048A 10.52 7.750e-14 923-937 PR00048A 10.52 1.643e-13 895-909 PR00048A 10.52 4.214e-13 1091-1105 PR00048A 10.52 4.857e-13 334-348 PR00048A 10.52 4.706e-12 362-376 PR00048B 6.02 5.000e-12 434-444 PR00048A 10.52 6.294e-12 474-488 PR00048A 10.52 9.471e-12 951-965 PR00048A 10.52 3.842e-11 979-993 PR00048A 10.52 4.789e-11 418-432 PR00048B 6.02 9.308e-11 1023-1033 PR00048B 6.02 1.563e-10 294-304 PR00048B 6.02 1.563e-10 883-893 PR00048B 6.02 2.125e-10 911-921 PR00048B 6.02 2.688e-10 350-360 PR00048B 6.02 8.313e-10 939-949 PR00048B 6.02 9.438e-10 490-500 PR00048B 6.02 1.000e-09 995-1005 PR00048B 6.02 3.368e-09 462-472 PR00048A 10.52 3.880e-09 390-404 PR00048B 6.02 4.316e-09 1079-1089 PR00048A 10.52 4.600e-09 867-881
1517	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 2.800e-14 325-338 PD00066 13.92 4.600e-14 437-450 PD00066 13.92 5.200e-14 998-1011 PD00066 13.92 6.400e-14 493-506 PD00066 13.92 8.800e-14 942-955 PD00066 13.92 5.500e-13 1026-1039 PD00066 13.92 6.000e-13 970-983 PD00066 13.92 6.500e-13 297-310 PD00066 13.92 6.500e-13 886-899 PD00066 13.92 7.000e-13 465-478 PD00066 13.92

Table 3A  
348

SEQ ID	Database entry ID	Description	Results*
			7.000e-13 914-927 PD00066 13.92 9.000e-13 353-366 PD00066 13.92 9.500e-13 1082-1095 PD00066 13.92 4.808e-10 409-422 PD00066 13.92 4.000e-09 381-394 PD00066 13.92 5.800e-09 1054-1067
1517	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479A 19.86 7.319e-09 1025-1048
1518	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 7.158e-13 42-65 BL00290B 13.17 4.500e-12 98-116
1519	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354K 9.35 7.706e-28 71-103 DM01354L 11.73 8.250e-14 103-130
1521	BL00250	TGF-beta family proteins.	BL00250A 21.24 2.500e-31 640-676 BL00250B 27.37 2.000e-27 705-741
1521	BL01273	CoA transferases proteins.	BL01273C 12.54 1.000e-40 130-170 BL01273D 19.11 9.471e-28 206-250 BL01273B 14.85 9.830e-20 81-115 BL01273A 12.56 2.286e-16 62-75
1521	PR00669	INHIBIN ALPHA CHAIN SIGNATURE	PR00669E 16.28 7.078e-12 640-658
1523	BL00427	Disintegrins proteins.	BL00427 13.93 4.100e-15 455-510
1523	PR00289	DISINTEGRIN SIGNATURE	PR00289A 13.62 3.700e-18 469-489 PR00289B 11.79 3.143e-12 498-511
1523	PR00011	TYPE III EGF-LIKE SIGNATURE	PR00011D 14.03 9.870e-09 675-694
1524	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 8.683e-12 186-197
1524	PR00700	PROTEIN TYROSINE PHOSPHATASE SIGNATURE	PR00700D 12.47 3.100e-09 183-202
1525	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237G 19.63 7.000e-17 35-62
1525	BL00237	G-protein coupled receptors proteins.	BL00237D 11.23 3.500e-14 45-62
1525	PR00424	ADENOSINE RECEPTOR SIGNATURE	PR00424F 8.50 2.688e-12 35-46
1525	PR00554	ADENOSINE A2B RECEPTOR SIGNATURE	PR00554F 8.86 1.750e-16 24-37 PR00554G 14.18 7.288e-12 69-89
1528	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 9.217e-09 338-351
1530	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354N 13.17 1.000e-40 36-81 DM01354O 8.73 1.000e-40 87-133 DM01354Q 9.23 1.000e-40 167-222 DM01354T 9.48 1.000e-40 273-316 DM01354V 12.97 1.000e-40 382-429 DM01354Y 10.69 1.000e-40 498-538 DM01354X 13.86 5.219e-39 449-488 DM01354P 9.18 2.723e-37 133-167 DM01354R 8.50 7.395e-32 222-252 DM01354W 12.64 3.000e-25 429-449 DM01354S 11.61 9.550e-23 252-273 DM01354U 12.24 2.350e-20 362-382

Table 3A  
349

SEQ ID	Database entry ID	Description	Results*
1530	DM00895	7 kw REVERSE TRANSCRIPTASE RNA POLYMERASE.	DM00895E 15.72 3.889e-11 149-174
1532	PR00036	LACI BACTERIAL REGULATORY PROTEIN HTH SIGNATURE	PR00036B 10.57 1.000e-08 282-293
1533	BL01160	Kinesin light chain repeat proteins.	BL01160D 10.17 7.077e-09 218-247
1533	PD00126	PROTEIN REPEAT DOMAIN TPR NUCLEA.	PD00126A 22.53 1.000e-08 485-506
1535	PR00671	INHIBIN BETA B CHAIN SIGNATURE	PR00671C 4.18 9.690e-09 44-64
1538	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 9.217e-14 89-120 BL00107B 13.31 6.786e-11 159-175
1538	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 6.766e-09 89-108
1539	PR00502	MUTT DOMAIN SIGNATURE	PR00502B 15.92 7.000e-09 131-147
1540	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 1.786e-10 813-828
1541	DM01418	352 FIBRILLAR COLLAGEN CARBOXYL-TERMINAL.	DM01418A 20.83 1.000e-40 631-679 DM01418C 20.48 1.000e-40 800-842 DM01418B 22.51 4.103e-38 729-771
1541	BL00420	Speract receptor repeat proteins domain proteins.	BL00420A 20.42 2.180e-10 549-578
1541	BL01113	C1q domain proteins.	BL01113A 17.99 1.957e-10 555-582 BL01113A 17.99 2.149e-10 549-576 BL01113A 17.99 7.319e-10 552-579
1542	BL00279	Membrane attack complex components / perforin proteins.	BL00279E 37.11 9.053e-12 901-949
1542	PR00907	THROMBOMODULIN SIGNATURE	PR00907B 11.29 8.274e-11 1016-1033
1542	PR00764	COMPLEMENT C9 SIGNATURE	PR00764F 16.89 3.836e-10 918-939
1542	DM00864	EGF-LIKE DOMAIN.	DM00864B 11.34 6.910e-10 1025-1044
1542	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 4.150e-14 1020-1036 BL01187B 12.04 2.575e-09 927-943
1542	BL01177	Anaphylatoxin domain proteins.	BL01177E 20.64 7.632e-09 1018-1045
1542	BL00799	Granulins proteins.	BL00799B 11.02 9.679e-09 1011-1047
1542	PR00010	TYPE II EGF-LIKE SIGNATURE	PR00010A 11.79 5.154e-11 912-924 PR00010C 11.16 7.545e-10 1025-1036 PR00010C 11.16 9.786e-09 932-943
1543	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 7.038e-09 79-113
1544	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972A 11.93 6.318e-19 478-496 BL00972D 22.55 7.968e-16 762-787 BL00972B 9.45 1.600e-12 559-569
1544	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 8.008e-13 235-250 PR00049D 0.00 7.375e-12 239-254

Table 3A  
350

SEQ ID	Database entry ID	Description	Results*
			PR00049D 0.00 5.916e-11 242-257 PR00049D 0.00 6.748e-11 236-251 PR00049D 0.00 9.395e-11 240-255 PR00049D 0.00 1.286e-10 233-248 PR00049D 0.00 8.929e-10 241-256 PR00049D 0.00 2.678e-09 243-258 PR00049D 0.00 4.051e-09 237-252 PR00049D 0.00 4.051e-09 238-253 PR00049D 0.00 4.051e-09 244-259
1544	PR00211	GLUTELIN SIGNATURE	PR00211B 0.86 7.500e-09 238-259
1544	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 5.339e-10 222-255 DM00215 19.43 7.268e-10 226-259 DM00215 19.43 2.525e-09 220-253 DM00215 19.43 9.695e-09 221-254
1544	BL00048	Protamine P1 proteins.	BL00048 6.39 9.888e-09 259-286
1546	PR00416	EUKARYOTIC DNA TOPOISOMERASE I SIGNATURE	PR00416B 15.58 5.000e-09 389-409
1546	BL00048	Protamine P1 proteins.	BL00048 6.39 8.200e-09 510-537
1546	PR00138	MATRIXIN SIGNATURE	PR00138E 6.01 8.429e-09 146-160
1546	BL00024	Hemopexin domain proteins.	BL00024E 7.58 8.694e-09 146-160
1546	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 4.857e-12 317-331 PR00019A 11.19 8.333e-09 133-147 PR00019B 11.36 8.920e-09 269-283
1547	DM01595	kw ALLANTOICASE SPAC1F7.09C.	DM01595B 10.54 2.625e-15 143-171 DM01595I 8.91 2.714e-15 457-475 DM01595D 10.94 3.274e-11 456-481
1549	PR00192	F-ACTIN CAPPING PROTEIN BETA SUBUNIT SIGNATURE	PR00192A 8.23 1.474e-27 47-68
1549	BL00231	F-actin capping protein beta subunit proteins.	BL00231A 8.59 1.829e-24 47-93
1551	BL00708	Prolyl endopeptidase family serine proteins.	BL00708B 24.91 7.197e-12 439-470
1551	PF00930	Dipeptidyl peptidase IV (DPP IV) N-terminal region.	PF00930I 15.96 6.373e-17 481-509 PF00930H 20.16 2.482e-13 402-445 PF00930J 8.78 1.000e-11 533-554 PF00930G 21.30 9.613e-09 362-400
1552	BL00878	Orn/DAP/Arg decarboxylases family 2 pyridoxal-P attachment si.	BL00878F 19.67 7.231e-20 171-194 BL00878C 17.74 3.676e-11 38-58
1552	PR00163	RUBREDOXIN SIGNATURE	PR00163B 14.45 8.200e-09 143-160
1555	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354V 12.97 1.000e-40 705-752 DM01354U 12.24 6.700e-22 685-705 DM01354T 9.48 8.355e-15 596-639 DM01354W 12.64 9.122e-13 752-772
1557	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354N 13.17 1.000e-40 158-203 DM01354O 8.73 1.000e-40 209-255 DM01354Q 9.23 1.000e-40 289-344 DM01354T 9.48 1.000e-40 398-441 DM01354V 12.97 1.000e-40 507-554 DM01354P 9.18 5.213e-37 255-289

Table 3A  
351

SEQ ID	Database entry ID	Description	Results*
			DM01354R 8.50 6.760e-28 344-374 DM01354L 11.73 1.200e-26 101-128 DM01354S 11.61 7.300e-26 377-398 DM01354M 12.50 6.094e-24 128-158 DM01354U 12.24 6.700e-22 487-507 DM01354K 9.35 1.228e-17 69-101 DM01354W 12.64 1.000e-13 554-574
1557	DM00895	7 kw REVERSE TRANSCRIPTASE RNA POLYMERASE.	DM00895E 15.72 7.828e-10 271-296
1558	PR00834	HTRA/DEGQ PROTEASE FAMILY SIGNATURE	PR00834F 10.91 2.946e-09 194-207
1558	BL00741	Guanine-nucleotide dissociation stimulators CDC24 family sign.	BL00741B 14.27 2.962e-09 1104-1127
1558	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 4.814e-09 1318-1333 PR00049D 0.00 5.729e-09 259-274
1558	PR00554	ADENOSINE A2B RECEPTOR SIGNATURE	PR00554B 12.52 8.855e-09 917-926
1560	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 3.700e-10 977-1012
1565	BL01226	Hydroxymethylglutaryl-coenzyme A synthase proteins.	BL01226A 13.79 1.000e-40 50-89 BL01226C 13.51 1.000e-40 127-167 BL01226H 17.74 1.000e-40 478-526 BL01226G 15.76 3.483e-32 384-413 BL01226B 13.35 1.818e-31 95-127 BL01226F 9.78 8.714e-23 345-363 BL01226E 13.74 7.716e-22 304-345 BL01226D 11.60 5.680e-18 226-262 BL01226D 11.60 6.940e-10 174-210
1568	BL00189	2-oxo acid dehydrogenases acyltransferase component lipoyl bi.	BL00189A 19.58 4.414e-13 103-138
1574	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 4.661e-09 254-269
1574	PR00456	RIBOSOMAL PROTEIN P2 SIGNATURE	PR00456E 3.06 6.241e-09 184-199
1577	PR00756	MEMBRANE ALANYL DIPEPTIDASE (M1) FAMILY SIGNATURE	PR00756B 14.06 2.364e-13 232-248 PR00756C 11.60 7.300e-12 310-321
1579	BL00048	Protamine P1 proteins.	BL00048 6.39 3.605e-10 636-663 BL00048 6.39 5.500e-10 630-657 BL00048 6.39 1.225e-09 628-655 BL00048 6.39 4.263e-09 648-675 BL00048 6.39 5.388e-09 627-654 BL00048 6.39 8.988e-09 626-653
1580	BL00284	Serpins proteins.	BL00284C 28.56 5.821e-27 192-234 BL00284A 15.64 1.000e-18 63-87 BL00284D 16.34 1.500e-14 407-434 BL00284B 17.99 1.273e-12 162-183
1581	BL01052	Calponin family repeat proteins.	BL01052B 15.31 3.308e-11 110-136
1581	PR00888	SMOOTH MUSCLE	PR00888C 12.27 2.141e-09 110-

Table 3A  
352

SEQ ID	Database entry ID	Description	Results*
		PROTEIN/CALPONIN FAMILY SIGNATURE	126
1582	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 6.500e-11 140-154 PR00019B 11.36 7.750e-10 92-106 PR00019A 11.19 5.000e-09 95-109
1582	PR00423	CELL DIVISION PROTEIN FTSZ SIGNATURE	PR00423E 7.36 9.550e-09 505-527
1587	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 2.500e-10 277-294
1587	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 8.286e-10 186-195
1588	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354R 8.50 7.750e-29 192-222 DM01354S 11.61 7.300e-26 222-243
1589	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354U 12.24 5.891e-19 62-82 DM01354V 12.97 1.175e-09 82-129
1590	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354Z 9.06 5.065e-13 167-216
1591	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354R 8.50 1.840e-18 141-171
1591	BL00300	SRP54-type proteins GTP-binding domain proteins.	BL00300C 25.57 7.100e-09 648-702
1593	BL00573	Pyridine nucleotide-disulphide oxidoreductases class-II activ.	BL00573A 9.65 7.273e-10 750-768
1593	PR00411	PYRIDINE NUCLEOTIDE DISULPHIDE REDUCTASE CLASS-I SIGNATURE	PR00411A 15.95 4.569e-09 747-770
1593	PR00350	VITAMIN D RECEPTOR SIGNATURE	PR00350E 11.55 5.546e-09 13-33
1593	PR00368	FAD-DEPENDENT PYRIDINE NUCLEOTIDE REDUCTASE SIGNATURE	PR00368A 17.76 7.589e-09 747-770
1593	BL01113	C1q domain proteins.	BL01113A 17.99 9.827e-09 484-511
1593	PR00308	TYPE I ANTIFREEZE PROTEIN SIGNATURE	PR00308C 3.83 8.892e-10 373-383 PR00308C 3.83 1.701e-09 375-385 PR00308A 5.90 4.282e-09 752-767 PR00308A 5.90 6.824e-09 372-387 PR00308C 3.83 1.000e-08 324-334
1594	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354X 13.86 6.595e-36 13-52 DM01354Y 10.69 3.455e-09 62-102
1595	PR00756	MEMBRANE ALANYL DIPEPTIDASE (M1) FAMILY SIGNATURE	PR00756D 10.58 7.000e-19 187-203 PR00756E 11.91 3.700e-15 206-219 PR00756C 11.60 4.000e-12 151-162
1595	BL00142	Neutral zinc metallopeptidases, zinc-binding region proteins.	BL00142 8.38 2.286e-09 187-198
1596	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354I 15.55 1.000e-40 119-160 DM01354V 12.97 1.209e-14 401-448 DM01354H 18.00 5.280e-10 79-119
1600	BL00484	Thyroglobulin type-1 repeat proteins proteins.	BL00484C 17.01 3.854e-11 96-111



Table 3A  
353

SEQ ID	Database entry ID	Description	Results*
1601	PR00342	RHESUS BLOOD GROUP PROTEIN SIGNATURE	PR00342A 9.71 1.000e-23 12-31
1608	PD01941	TRANSMEMBRANE COTRANSPORTER SYMP.	PD01941E 15.92 1.000e-40 135-182 PD01941F 28.52 9.229e-17 254-309
1609	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354W 12.64 9.375e-16 123-143
1614	BL00420	Speract receptor repeat proteins domain proteins.	BL00420B 22.67 5.808e-26 49-104 BL00420C 11.90 2.500e-10 134-145
1614	PR00258	SPERACT RECEPTOR SIGNATURE	PR00258B 9.63 7.947e-11 64-76 PR00258C 9.05 3.204e-10 79-90 PR00258E 13.33 3.333e-10 133-146 PR00258D 14.41 3.286e-09 110-125
1616	BL00326	Tropomyosins proteins.	BL00326D 8.76 1.505e-09 377-418
1616	PR00261	LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR SIGNATURE	PR00261D 12.47 3.338e-10 43-65 PR00261E 11.08 4.575e-10 84-106 PR00261C 11.37 5.050e-10 43-65 PR00261F 11.57 5.920e-10 43-65 PR00261A 11.02 6.769e-10 43-65 PR00261A 11.02 9.192e-10 84-106 PR00261D 12.47 2.667e-09 84-106 PR00261F 11.57 4.304e-09 84-106 PR00261C 11.37 5.500e-09 84-106 PR00261B 14.12 5.714e-09 84-106 PR00261B 14.12 6.464e-09 43-65
1616	BL01209	LDL-receptor class A (LDLRA) domain proteins.	BL01209 9.31 6.464e-09 93-106
1616	BL00412	Neuromodulin (GAP-43) proteins.	BL00412D 16.54 5.613e-13 306-357 BL00412D 16.54 8.821e-12 301-352 BL00412D 16.54 1.098e-10 307-358 BL00412D 16.54 9.609e-10 304-355 BL00412D 16.54 1.184e-09 308-359 BL00412D 16.54 1.918e-09 309-360 BL00412D 16.54 2.745e-09 302-353 BL00412D 16.54 5.684e-09 299-350 BL00412D 16.54 8.347e-09 303-354
1616	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 8.640e-09 319-354
1616	BL00422	Granins proteins.	BL00422C 16.18 8.085e-10 326-354 BL00422C 16.18 8.468e-10 330-358 BL00422C 16.18 8.851e-10 323-351 BL00422C 16.18 4.529e-09 325-353 BL00422C 16.18 9.471e-09 318-346
1618	PR00681	RIBOSOMAL PROTEIN S1 SIGNATURE	PR00681I 8.81 9.897e-09 412-431
1619	BL00134	Serine proteases, trypsin family, histidine proteins.	BL00134A 11.96 3.813e-15 55-72
1619	PR00722	CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1) SIGNATURE	PR00722A 12.27 8.448e-14 56-72
1619	BL00021	Kringle domain proteins.	BL00021B 13.33 1.763e-13 55-73

Table 3A  
354

SEQ ID	Database entry ID	Description	Results*
1619	PR00839	V8 SERINE PROTEASE FAMILY SIGNATURE	PR00839B 11.20 4.945e-09 55-73
1620	BL00125	Serine/threonine specific protein phosphatases proteins.	BL00125C 19.97 1.000e-40 86-133 BL00125D 33.11 9.679e-38 154-209
1620	PR00114	SERINE/THREONINE PHOSPHATASE FAMILY SIGNATURE	PR00114E 17.75 2.500e-27 106-134 PR00114D 12.34 9.591e-26 77-104 PR00114F 17.51 3.400e-17 170-191 PR00114G 17.20 3.676e-17 198-215
1625	BL01103	Aspartate-semialdehyde dehydrogenase proteins.	BL01103F 16.78 3.625e-24 256-282 BL01103C 15.07 6.727e-20 91-113 BL01103B 20.29 5.846e-16 51-73 BL01103E 11.30 4.115e-14 238-251 BL01103D 7.29 1.000e-09 155-166
1626	PD02410	SYNTHETASE BETA CHAIN LIGASE PHENYLALANYL-TRNA.	PD02410E 10.29 2.241e-13 83-97
1628	BL00044	Bacterial regulatory proteins, lysR family proteins.	BL00044 19.03 7.188e-11 21-55
1631	BL00370	PEP-utilizing enzymes phosphorylation site proteins proteins.	BL00370C 11.73 9.182e-17 96-110
1632	PR00101	ASPARTATE CARBAMOYLTRANSFERASE SIGNATURE	PR00101C 12.35 9.328e-10 410-428
1632	BL00097	Aspartate and ornithine carbamoyltransferases proteins.	BL00097C 16.87 1.621e-09 411-425
1634	BL00530	Ribonuclease T2 family histidine proteins 1.	BL00530C 11.43 1.000e-17 101-113
1643	BL00050	Ribosomal protein L23 proteins.	BL00050A 23.71 3.659e-18 217-250
1645	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 5.571e-14 479-507 PD01719A 12.89 2.000e-09 666-694 PD01719A 12.89 4.889e-09 606-634 PD01719A 12.89 8.111e-09 419-447
1647	BL00564	Argininosuccinate synthase proteins.	BL00564D 22.43 2.723e-22 157-194 BL00564A 19.93 6.582e-11 37-74
1651	PR00453	VON WILLEBRAND FACTOR TYPE A DOMAIN SIGNATURE	PR00453A 12.79 9.719e-13 88-106 PR00453B 14.65 1.818e-12 125-140 PR00453C 12.26 3.769e-10 246-255
1652	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 9.471e-27 84-118 DM00031A 16.80 7.055e-25 206-254 DM00031A 16.80 5.368e-18 20-68 DM00031B 15.41 8.519e-09 267-301
1653	PR00173	GLUTAMATE-ASPARTATE SYMPORTER SIGNATURE	PR00173F 10.44 8.967e-09 8-28
1654	BL01032	Protein phosphatase 2C proteins.	BL01032C 6.14 1.000e-10 118-128
1657	BL00523	Sulfatases proteins.	BL00523A 13.36 2.200e-13 47-64
1658	BL00194	Thioredoxin family proteins.	BL00194 12.16 4.857e-16 482-495 BL00194 12.16 3.813e-15 347-360
1658	PR00421	THIOREDOXIN FAMILY SIGNATURE	PR00421B 11.40 3.348e-12 489-499 PR00421B 11.40 9.379e-11

Table 3A  
355

SEQ ID	Database entry ID	Description	Results*
			354-364 PR00421A 10.15 4.000e-10 346-355 PR00421A 10.15 4.857e-10 481-490 PR00421C 13.60 5.714e-10 533-545
1662	PF00850	Histone deacetylase family.	PF00850E 8.88 1.766e-21 1269-1295 PF00850F 15.70 1.889e-10 1307-1340 PF00850G 22.75 3.724e-10 1346-1388
1666	BL00317	WAP-type 'four-disulfide core' domain proteins.	BL00317B 14.58 7.231e-22 92-114
1666	PR00003	4-DISULPHIDE CORE SIGNATURE	PR00003C 7.69 6.885e-11 98-108 PR00003D 8.10 4.971e-09 108-117
1666	BL00313	Seminal vesicle protein I repeat proteins proteins.	BL00313B 8.87 1.000e-40 31-68 BL00313C 11.19 1.000e-18 103-118 BL00313B 8.87 5.337e-17 25-62 BL00313B 8.87 7.593e-16 19-56 BL00313B 8.87 1.910e-15 43-80 BL00313B 8.87 9.413e-14 37-74 BL00313B 8.87 4.522e-09 49-86 BL00313A 10.60 5.500e-09 1-9 BL00313C 11.19 9.060e-09 30-45
1667	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 5.395e-31 108-147
1667	BL01277	Ribonuclease PH proteins.	BL01277A 17.39 8.818e-09 64-102
1672	DM00372	CARCINOEMBRYONIC ANTIGEN PRECURSOR AMINO-TERMINAL DOMAIN.	DM00372B 20.31 8.920e-15 363-408 DM00372B 20.31 3.329e-12 68-113
1673	PF00992	Troponin.	PF00992A 16.67 8.859e-10 176-211
1673	BL00093	N-4 cytosine-specific DNA methylases proteins.	BL00093A 9.77 4.329e-09 718-733
1673	PR00948	ELICITIN SIGNATURE	PR00948B 15.44 5.361e-08 707-732
1673	PR00558	ALPHA-2A ADRENERGIC RECEPTOR SIGNATURE	PR00558C 6.54 6.719e-08 429-449
1676	BL00962	Ribosomal protein S2 proteins.	BL00962B 36.15 6.529e-14 53-107
1678	PD00015	GLYCOPROTEIN PRECURSOR CELL SI.	PD00015A 8.90 6.400e-09 35-43
1678	PR00795	RYANODINE RECEPTOR SIGNATURE	PR00795C 6.30 8.252e-09 107-132
1683	BL00262	Insulin family proteins.	BL00262A 12.48 3.829e-09 30-48
1684	BL00061	Short-chain dehydrogenases/reductases family proteins.	BL00061B 25.79 8.105e-19 133-171
1684	PR00080	ALCOHOL DEHYDROGENASE SUPERFAMILY SIGNATURE	PR00080C 17.16 2.125e-10 153-173
1684	PR00081	GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE	PR00081A 10.53 7.805e-11 5-23 PR00081C 15.13 1.771e-09 127-144
1685	BL00983	Ly-6 / u-PAR domain proteins.	BL00983C 12.69 6.885e-09 197-213
1686	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 7.158e-18 678-709
1686	PR00653	ACTIVIN TYPE II RECEPTOR SIGNATURE	PR00653E 15.68 5.114e-16 871-891 PR00653D 13.25 7.195e-12 697-719
1686	BL00790	Receptor tyrosine kinase class V	BL00790N 13.25 7.505e-09 691-

Table 3A  
356

SEQ ID	Database entry ID	Description	Results*
		proteins.	718
1686	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 8.031e-09 678-697
1687	BL00649	G-protein coupled receptors family 2 proteins.	BL00649A 24.05 2.333e-22 85-113
1687	PR00491	VASOACTIVE INTESTINAL PEPTIDE RECEPTOR SIGNATURE	PR00491A 11.52 9.509e-10 109-121
1688	BL50002	Src homology 3 (SH3) domain proteins profile.	BL50002A 14.19 2.500e-18 154-173
1688	PR00401	SH2 DOMAIN SIGNATURE	PR00401A 14.00 5.765e-15 214-229
1688	BL50001	Src homology 2 (SH2) domain proteins profile.	BL50001A 10.81 6.250e-11 154-167
1688	PR00452	SH3 DOMAIN SIGNATURE	PR00452B 11.65 5.765e-14 164-180 PR00452A 10.48 8.500e-12 150-161 PR00452C 10.78 1.000e-10 181-191
1688	PR00499	NEUTROPHIL CYTOSOL FACTOR 2 SIGNATURE	PR00499D 10.18 7.279e-09 152-173
1693	DM01551	kw OSTEOINDUCTIVE YOPM MEMBRANE OUTER.	DM01551A 15.63 7.265e-09 559-589
1693	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 2.957e-10 614-628 PR00019B 11.36 4.600e-10 527-541 PR00019B 11.36 1.720e-09 611-625 PR00019A 11.19 3.000e-09 660-674 PR00019A 11.19 7.667e-09 530-544
1698	PR00528	GLUCOCORTICOID RECEPTOR SIGNATURE	PR00528C 14.75 9.438e-09 210-231
1705	BL00733	Ribosomal protein S26e proteins.	BL00733B 12.04 4.661e-25 165-198
1706	PD00078	REPEAT PROTEIN ANK NUCLEAR ANKYR.	PD00078B 13.14 6.500e-11 412-425
1706	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 7.000e-11 773-787 PR00019A 11.19 7.000e-09 776-790
1706	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 6.242e-13 386-441 PF00791C 20.98 7.088e-09 436-475 PF00791B 28.49 7.936e-09 455-510
1706	PF00023	Ank repeat proteins.	PF00023A 16.03 1.000e-11 419-435 PF00023B 14.20 9.591e-09 415-425
1708	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 3.000e-13 251-264 PD00066 13.92 3.571e-12 223-236
1708	PR00498	NEUTROPHIL CYTOSOL FACTOR 1 SIGNATURE	PR00498A 12.01 5.765e-09 353-362
1708	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048B 6.02 1.563e-10 220-230 PR00048A 10.52 2.565e-10 232-246 PR00048A 10.52 5.320e-09 50-64 PR00048A 10.52 8.200e-09 204-218
1708	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 5.154e-11 235-252 BL00028 16.07 3.400e-10 207-224 BL00028 16.07 9.743e-09 53-70
1711	PF00550	Phosphopantetheine attachment site	PF00550C 13.05 7.188e-10 235-

Table 3A  
357

SEQ ID	Database entry ID	Description	Results*
		proteins.	252
1712	BL00484	Thyroglobulin type-1 repeat proteins.	BL00484B 9.04 4.400e-12 289-303
1712	BL00222	Insulin-like growth factor binding proteins.	BL00222B 11.09 7.300e-17 65-81 BL00222A 11.34 9.438e-11 45-56
1715	PF00075	RNase H.	PF00075C 11.58 6.786e-13 60-72
1717	BL00477	Alpha-2-macroglobulin family thiolester region proteins.	BL00477G 19.43 3.739e-34 984-1016 BL00477A 13.50 6.000e-32 174-203 BL00477J 19.04 9.625e-31 1301-1332 BL00477F 17.34 1.931e-30 840-870 BL00477L 23.51 1.964e-30 1498-1531 BL00477K 17.42 6.400e-24 1443-1467 BL00477E 17.53 7.000e-22 810-831 BL00477I 18.76 6.750e-20 1088-1115 BL00477C 15.70 6.667e-17 291-308 BL00477B 9.05 7.207e-15 264-277 BL00477H 9.07 1.333e-14 1057-1069 BL00477D 12.73 2.059e-14 784-794
1717	BL01177	Anaphylatoxin domain proteins.	BL01177E 20.64 9.438e-12 845-872
1718	BL00232	Cadherins extracellular repeat proteins domain proteins.	BL00232B 32.79 8.594e-35 151-199 BL00232B 32.79 5.579e-22 417-465 BL00232A 27.72 1.000e-20 57-90 BL00232C 10.65 3.613e-14 415-433 BL00232B 32.79 4.872e-11 534-582
1718	PR00205	CADHERIN SIGNATURE	PR00205B 11.39 4.545e-15 415-433 PR00205A 14.73 5.600e-09 183-199
1718	BL00422	Granins proteins.	BL00422C 16.18 3.647e-09 226-254 BL00422C 16.18 9.294e-09 220-248
1718	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 2.256e-09 221-256 PF01140D 15.54 3.198e-09 217-252 PF01140D 15.54 9.791e-09 220-255
1722	BL00086	Cytochrome P450 cysteine heme-iron ligand proteins.	BL00086 20.87 8.714e-20 469-501
1722	PR00463	E-CLASS P450 GROUP I SIGNATURE	PR00463H 12.41 5.821e-10 469-480
1722	PR00408	MITOCHONDRIAL P450 SIGNATURE	PR00408F 11.33 6.000e-10 388-407
1722	PR00464	E-CLASS P450 GROUP II SIGNATURE	PR00464I 14.64 1.771e-13 479-503 PR00464E 18.28 9.617e-11 389-410 PR00464C 18.84 1.000e-10 330-359 PR00464H 13.32 3.596e-09 466-480
1722	PR00465	E-CLASS P450 GROUP IV SIGNATURE	PR00465C 21.23 5.125e-25 332-359 PR00465H 17.76 5.500e-22 479-498 PR00465G 13.06 1.281e-17 463-480 PR00465D 14.64 2.607e-17 390-407 PR00465F 13.37 6.000e-15 439-458 PR00465E 15.03 9.217e-15 423-

Table 3A  
358

SEQ ID	Database entry ID	Description	Results*
			438 PR00465B 16.84 4.908e-09 113-137
1722	PR00385	P450 SUPERFAMILY SIGNATURE	PR00385D 13.11 1.667e-09 470-480 PR00385C 16.94 4.682e-09 395-407 PR00385A 14.97 5.865e-09 341-359 PR00385E 12.66 7.000e-09 479-491
1725	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 7.158e-10 107-117
1727	PR00220	SYNAPTOPHYSIN/SYNAPTOPORIN FAMILY SIGNATURE	PR00220B 15.48 2.800e-31 47-73 PR00220C 11.05 9.000e-31 100-125 PR00220D 8.32 3.842e-30 132-156 PR00220A 10.93 1.857e-28 23-46 PR00220E 3.46 4.273e-24 199-218
1727	BL00604	Synaptophysin / synaptoporin proteins.	BL00604A 9.13 1.000e-40 26-81 BL00604C 14.66 1.000e-40 117-149 BL00604D 12.28 1.000e-40 149-184 BL00604E 8.32 1.000e-40 184-226 BL00604F 5.96 1.000e-40 231-276 BL00604B 9.95 2.895e-37 87-117 BL00604F 5.96 1.122e-09 263-308
1727	PR00761	BINDIN PRECURSOR SIGNATURE	PR00761A 5.81 3.662e-09 287-304
1727	BL00291	Prion protein.	BL00291A 4.49 8.138e-09 236-271
1728	BL00795	Involucrin proteins.	BL00795C 17.06 6.700e-09 159-204
1729	PR00747	GLYCOSYL HYDROLASE FAMILY 47 SIGNATURE	PR00747E 15.13 8.269e-18 225-243 PR00747H 12.76 5.610e-17 326-347 PR00747C 12.06 4.797e-11 145-164
1730	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209C 4.56 3.893e-09 77-91
1731	PR00244	NEUROKININ RECEPTOR SIGNATURE	PR00244H 13.12 9.357e-26 294-313 PR00244G 11.10 2.286e-15 268-280 PR00244F 10.47 3.132e-15 204-218 PR00244A 10.34 3.368e-15 29-41 PR00244B 12.45 1.000e-13 55-65 PR00244D 6.84 1.964e-13 125-136 PR00244C 13.44 3.118e-12 108-119 PR00244E 5.93 8.875e-12 183-193
1731	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237G 19.63 9.053e-19 275-302 PR00237F 13.57 5.800e-18 234-259 PR00237B 13.50 9.400e-17 68-90 PR00237A 11.48 5.765e-14 35-60 PR00237D 8.94 2.350e-11 133-155 PR00237E 13.03 1.600e-10 184-208
1731	BL00237	G-protein coupled receptors proteins.	BL00237C 13.19 2.688e-19 229-256 BL00237D 11.23 4.429e-13 285-302 BL00237B 5.28 5.909e-09 192-204
1732	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870D 15.74 8.755e-09 223-258
1733	BL00523	Sulfatases proteins.	BL00523A 13.36 8.800e-13 44-61

Table 3A  
359

SEQ ID	Database entry ID	Description	Results*
1734	PR00464	E-CLASS P450 GROUP II SIGNATURE	PR00464A 20.47 9.719e-13 167-188
1735	BL00414	Profilin proteins.	BL00414D 15.59 5.909e-10 1037-1064
1735	PR00392	PROFILIN SIGNATURE	PR00392C 18.98 5.592e-10 986-1007 PR00392D 12.00 7.652e-09 1008-1023
1738	PR00003	4-DISULPHIDE CORE SIGNATURE	PR00003C 7.69 9.357e-09 73-83
1741	PR00915	LUTEOVIRUS GROUP 1 COAT PROTEIN SIGNATURE	PR00915G 15.24 6.875e-10 303-325
1745	PR00837	ALLERGEN V5/TPX-1 FAMILY SIGNATURE	PR00837C 17.21 6.143e-20 126-143 PR00837A 14.77 1.973e-13 57-76
1745	BL01009	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 proteins.	BL01009D 14.19 4.300e-20 127-148 BL01009A 13.75 6.586e-13 57-75
1745	PR00838	VENOM ALLERGEN 5 SIGNATURE	PR00838G 16.07 2.033e-17 125-145 PR00838D 8.73 4.214e-09 57-76
1748	PR00920	SPUMAVIRUS ASPARTIC PROTEASE (A9) SIGNATURE	PR00920C 13.24 7.310e-09 212-234
1749	PR00672	INHIBIN BETA C CHAIN SIGNATURE	PR00672D 10.52 6.400e-09 96-113
1749	PD01572	PHOTOSYSTEM II REACTION CENTRE T PROTEIN PHOTOS.	PD01572 8.77 6.917e-09 544-574
1751	BL00713	Sodium:dicarboxylate symporter family proteins.	BL00713A 20.35 2.227e-35 314-356 BL00713B 13.04 8.000e-22 375-395
1751	PR00173	GLUTAMATE-ASPARTATE SYMPORTER SIGNATURE	PR00173C 13.89 7.353e-25 327-353 PR00173B 9.62 9.053e-13 19-40
1752	PR00219	SYNAPTOBREVIN SIGNATURE	PR00219C 9.04 7.750e-09 1-21
1753	PR00379	INTEIN SIGNATURE	PR00379A 16.17 9.780e-09 393-407
1757	PF00783	Inositol polyphosphate phosphatase, catalytic domain proteins homologue.	PF00783A 11.91 3.400e-14 539-549
1758	PR00169	POTASSIUM CHANNEL SIGNATURE	PR00169B 14.43 2.800e-30 239-268 PR00169C 16.31 3.109e-11 307-331
1762	BL00456	Sodium:solute symporter family proteins.	BL00456A 22.59 1.000e-40 16-71 BL00456C 24.55 1.000e-40 154-209 BL00456B 18.94 7.158e-28 92-122 BL00456D 6.92 2.862e-09 450-460
1763	DM01253	BNR MOTIF REPEAT.	DM01253A 11.72 7.785e-09 454-470
1763	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 2.200e-12 86-100 PR00019B 11.36 6.400e-10 83-97 PR00019A 11.19 2.333e-09 278-292 PR00019B 11.36 4.960e-09 275-289 PR00019A 11.19 9.333e-09 62-76
1764	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 5.714e-14 421-449
1765	PR00824	HEPATIC LIPASE SIGNATURE	PR00824A 7.81 7.214e-22 6-25
1768	BL00979	G-protein coupled receptors family 3	BL00979A 19.66 1.000e-40 77-125

Table 3A  
360

SEQ ID	Database entry ID	Description	Results*
		proteins.	
1769	PF00168	C2 domain proteins.	PF00168C 27.49 4.000e-12 410-436
1769	PR00360	C2 DOMAIN SIGNATURE	PR00360B 13.61 5.345e-12 292-306 PR00360B 13.61 1.000e-11 427-441 PR00360A 14.59 2.895e-09 398-411
1769	PR00399	SYNAPTOTAGMIN SIGNATURE	PR00399B 14.27 8.425e-12 385-399 PR00399D 14.48 7.796e-09 328-339
1771	BL00312	Glycophorin A proteins.	BL00312B 9.22 5.781e-10 717-746
1771	BL00446	RNA polymerases D / 30 to 40 Kd subunits proteins.	BL00446C 18.90 9.087e-09 1326-1353
1771	PR00122	VACUOLAR ATP SYNTHASE 16 KD SUBUNIT SIGNATURE	PR00122D 9.97 9.330e-09 716-740
1771	PR00205	CADHERIN SIGNATURE	PR00205B 11.39 5.745e-10 225-243 PR00205B 11.39 4.966e-09 333-351 PR00205B 11.39 9.390e-09 441-459
1771	BL00232	Cadherins extracellular repeat proteins domain proteins.	BL00232B 32.79 9.308e-18 227-275 BL00232B 32.79 9.206e-17 443-491 BL00232B 32.79 3.407e-11 118-166 BL00232C 10.65 4.115e-10 225-243 BL00232C 10.65 4.462e-10 546-564 BL00232C 10.65 7.404e-10 333-351 BL00232C 10.65 9.842e-09 441-459
1773	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354Z 9.06 1.000e-16 102-151
1774	PR00259	TRANSMEMBRANE FOUR FAMILY SIGNATURE	PR00259C 16.40 5.313e-17 513-542 PR00259B 14.81 4.000e-14 486-513 PR00259D 13.50 2.340e-13 676-703
1774	BL00421	Transmembrane 4 family proteins.	BL00421B 17.62 1.000e-27 492-531 BL00421E 20.97 6.211e-13 673-703
1775	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245C 7.84 8.258e-17 319-335 PR00245A 18.03 9.265e-16 52-74 PR00245B 10.38 6.143e-14 258-273
1775	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 5.974e-12 83-123 BL00237A 27.68 5.974e-12 171-211
1775	PR00534	MELANOCORTIN RECEPTOR FAMILY SIGNATURE	PR00534A 11.49 6.123e-10 44-57
1775	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237E 13.03 9.333e-11 280-304 PR00237A 11.48 5.935e-09 19-44
1778	BL00417	Synaptobrevin proteins.	BL00417B 18.48 2.414e-19 141-195 BL00417A 7.74 6.704e-13 113-141
1778	PR00219	SYNAPTOSOMAL SIGNATURE	PR00219A 8.98 9.156e-13 118-138
1780	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290B 13.17 2.385e-15 255-273 BL00290A 20.89 1.529e-14 34-57 BL00290A 20.89 6.684e-13



Table 3A  
361

SEQ ID	Database entry ID	Description	Results*
			198-221
1782	BL00211	ABC transporters family proteins.	BL00211B 13.37 2.385e-17 558-590
1782	BL00674	AAA-protein family proteins.	BL00674B 4.46 5.622e-09 448-470
1782	PR00671	INHIBIN BETA B CHAIN SIGNATURE	PR00671B 4.29 9.408e-09 396-416
1782	PR00300	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT SIGNATURE	PR00300A 9.56 9.649e-09 451-470
1783	PR00176	SODIUM/NEUROTRANSMITTER SYMPORTER SIGNATURE	PR00176C 10.84 8.615e-25 283-310 PR00176A 16.82 5.909e-24 211-233 PR00176B 7.31 4.600e-23 240-260 PR00176D 9.02 6.143e-21 412-430 PR00176F 10.73 1.333e-20 548-568 PR00176E 11.41 2.227e-20 494-515 PR00176G 12.48 5.034e-17 630-651 PR00176H 15.27 7.429e-17 670-691
1783	BL00610	Sodium:neurotransmitter symporter family proteins.	BL00610A 17.73 1.000e-40 211-261 BL00610B 23.65 1.000e-40 274-324 BL00610D 20.97 1.000e-40 451-504 BL00610F 29.02 1.000e-40 641-696 BL00610E 20.34 9.100e-37 544-587 BL00610G 12.89 6.087e-22 700-723 BL00610C 12.94 4.282e-15 385-437
1784	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 7.592e-09 799-842
1785	BL00349	CTF/NF-I proteins.	BL00349A 10.07 1.000e-40 649-695 BL00349C 9.33 1.000e-40 723-766 BL00349E 10.79 9.833e-36 793-836 BL00349B 10.51 1.205e-34 695-723 BL00349D 11.70 1.000e-33 766-793
1786	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 9.100e-11 243-257 PR00019B 11.36 9.100e-10 431-445 PR00019A 11.19 2.000e-09 434-448 PR00019A 11.19 4.667e-09 482-496 PR00019B 11.36 6.040e-09 643-657 PR00019B 11.36 6.400e-09 597-611 PR00019B 11.36 9.640e-09 479-493
1789	PR00702	ACRIFLAVIN RESISTANCE PROTEIN FAMILY SIGNATURE	PR00702G 9.63 4.214e-22 74-98 PR00702F 12.14 1.474e-19 49-73 PR00702H 12.94 7.955e-17 153-171 PR00702I 17.85 7.857e-16 216-231
1790	DM00914	BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEMS INNER MEMBRANE.	DM00914C 21.08 8.457e-35 121-168 DM00914B 22.92 5.750e-30 45-86
1790	PR00797	STREPTOPAIN (C10) CYSTEINE PROTEASE FAMILY SIGNATURE	PR00797F 12.40 9.804e-09 41-63
1791	BL00942	glpT family of transporters proteins.	BL00942E 21.14 9.700e-40 104-

Table 3A  
362

SEQ ID	Database entry ID	Description	Results*
			145 BL00942D 18.76 1.273e-31 55-92 BL00942C 14.04 2.688e-21 13-33 BL00942F 15.07 7.632e-21 162-180
1792	BL00523	Sulfatases proteins.	BL00523C 12.64 5.800e-12 593-604
1796	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 7.261e-10 315-329 PR00019B 11.36 4.600e-09 312-326
1798	PR00303	PREPROTEIN TRANSLOCASE SECY SUBUNIT SIGNATURE	PR00303E 14.69 1.429e-22 98-122 PR00303D 16.15 8.650e-22 72-98 PR00303C 15.93 9.036e-19 34-58
1798	BL00755	Protein secY proteins.	BL00755C 14.21 2.500e-18 89-105
1798	PR00825	VESPID VENOM ALLERGEN PHOSPHOLIPASE A1 SIGNATURE	PR00825B 14.81 9.738e-09 234-255
1800	PR00081	GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE	PR00081A 10.53 1.439e-11 179-197
1800	PR00669	INHIBIN ALPHA CHAIN SIGNATURE	PR00669B 8.27 2.976e-09 80-97
1801	BL01022	PTR2 family proton/oligopeptide symporters proteins.	BL01022E 23.51 1.173e-12 474-510 BL01022D 9.42 3.455e-11 12-25
1802	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 2.350e-14 489-505
1802	PR00792	PEPSIN (A1) ASPARTIC PROTEASE FAMILY SIGNATURE	PR00792B 12.78 7.750e-14 331-345 PR00792C 9.10 1.000e-12 380-392
1802	BL00141	Eukaryotic and viral aspartyl proteases proteins.	BL00141C 9.74 1.000e-11 336-346 BL00141D 6.28 3.700e-11 380-390
1802	BL01177	Anaphylatoxin domain proteins.	BL01177C 17.39 8.780e-10 483-502
1804	PR00165	ANION EXCHANGER SIGNATURE	PR00165F 10.39 9.667e-12 32-51
1804	BL00219	Anion exchangers family proteins.	BL00219L 18.71 1.000e-40 126-165 BL00219N 10.66 6.164e-31 194-238 BL00219M 9.98 3.100e-17 148-194 BL00219P 19.59 6.123e-13 247-274 BL00219H 10.06 4.555e-11 11-59 BL00219K 12.73 8.780e-10 84-126
1805	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972A 11.93 3.118e-16 786-804 BL00972D 22.55 1.500e-14 1281-1306 BL00972E 20.72 6.850e-13 1308-1330 BL00972B 9.45 7.923e-10 882-892
1810	BL00216	Sugar transport proteins.	BL00216B 27.64 8.531e-13 736-786
1810	DM00973	3 kw RESISTANCE BENOMYL YLL028W CYCLOHEXIMIDE.	DM00973A 21.17 1.391e-11 708-745
1813	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 9.308e-15 288-301
1813	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048B 6.02 7.188e-10 285-295 PR00048A 10.52 8.043e-10 269-283 PR00048A 10.52 3.880e-09 297-311
1813	BL00028	Zinc finger, C2H2 type, domain	BL00028 16.07 1.692e-11 515-532

Table 3A  
363

SEQ ID	Database entry ID	Description	Results*
		proteins.	BL00028 16.07 4.000e-10 487-504 BL00028 16.07 1.257e-09 773-790 BL00028 16.07 5.371e-09 272-289
1813	PF01059	NADH-ubiquinone oxidoreductase chain 4, amino terminus.	PF01059A 13.84 6.192e-09 19-53
1815	BL00134	Serine proteases, trypsin family, histidine proteins.	BL00134B 15.99 8.560e-17 316-340 BL00134A 11.96 4.214e-16 53-70
1815	BL00495	Apple domain proteins.	BL00495M 8.50 5.071e-16 226-261 BL00495N 11.04 4.438e-15 308-343
1815	PR00722	CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1) SIGNATURE	PR00722A 12.27 4.414e-14 54-70 PR00722C 10.87 9.471e-13 315-328
1815	BL00021	Kringle domain proteins.	BL00021B 13.33 6.580e-16 53-71 BL00021C 22.21 2.125e-10 235-257
1815	BL01253	Type I fibronectin domain proteins.	BL01253G 11.34 4.316e-12 315-329 BL01253E 16.01 9.899e-11 224-261 BL01253D 4.84 5.920e-09 53-67
1815	PR00839	V8 SERINE PROTEASE FAMILY SIGNATURE	PR00839B 11.20 6.425e-09 53-71
1816	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354H 18.00 5.618e-12 134-174
1817	BL00859	GTP cyclohydrolase I proteins.	BL00859B 13.15 9.057e-09 72-113
1819	PR00700	PROTEIN TYROSINE PHOSPHATASE SIGNATURE	PR00700D 12.47 2.421e-20 1431-1450 PR00700C 13.17 5.765e-17 1390-1408 PR00700F 11.18 7.429e-12 1478-1489
1819	PR00213	MYELIN P0 PROTEIN SIGNATURE	PR00213E 5.51 1.656e-09 1186-1211
1819	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 1.000e-14 1434-1445 BL00383D 11.92 1.692e-14 1394-1407 BL00383A 13.34 6.850e-14 1280-1295 BL00383F 15.51 4.240e-13 1472-1488 BL00383C 10.10 6.625e-10 1318-1329 BL00383B 7.61 4.000e-09 1305-1314
1819	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014C 15.44 1.450e-10 688-707 PR00014C 15.44 6.400e-10 866-885 PR00014C 15.44 2.565e-09 510-529 PR00014C 15.44 5.696e-09 1056-1075
1821	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245C 7.84 4.073e-15 361-377 PR00245E 12.40 8.286e-12 414-429
1821	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237G 19.63 5.814e-09 395-422
1822	PF00242	DNA polymerase (viral) N-terminal domain proteins.	PF00242F 12.18 8.522e-09 322-344

\* Results include in order: accession number subtype; raw score; p-value; position of signature in amino acid sequence

Table 3B  
364

SEQ ID NO:	Database entry ID	Description	Results*
912	IPB001398A	Macrophage migration inhibitory factor family	23.60 6.73e-29 124-163
912	IPB001398B	Macrophage migration inhibitory factor family	19.16 2.06e-20 164-190
912	IPB001398C	Macrophage migration inhibitory factor family	26.46 6.96e-11 192-238
913	IPB000033A	Low-density lipoprotein (ldl) receptor, YWTD repeat	21.82 4.35e-09 514-536
913	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 4.48e-10 817-827
913	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 7.10e-10 574-584
913	IPB000118B	Granulin	7.94 9.20e-09 472-510
913	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 1.96e-11 484-499
913	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 3.84e-14 696-711
913	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 3.86e-13 569-584
913	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 6.54e-17 812-827
913	IPB000561	EGF-like domain	4.89 3.57e-09 821-829
913	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 8.02e-10 175-190
913	IPB001881B	Calcium-binding EGF-like domain	12.28 4.79e-11 696-707
913	IPB001881B	Calcium-binding EGF-like domain	12.28 5.20e-12 812-823
913	IPB001881B	Calcium-binding EGF-like domain	12.28 8.58e-11 527-538
913	IPB001881B	Calcium-binding EGF-like domain	12.28 8.71e-10 569-580
913	IPB001881B	Calcium-binding EGF-like domain	12.28 9.53e-11 484-495
913	IPB002861B	Reeler domain	10.50 4.48e-09 124-152
913	IPB002861B	Reeler domain	10.50 8.98e-09 181-209
913	IPB002861C	Reeler domain	23.17 6.48e-09 111-165
913	IPB003367A	Thrombospondin type 3 repeat	11.78 5.13e-09 782-802
913	IPB003367A	Thrombospondin type 3 repeat	11.78 5.83e-10 577-597
913	IPB003886D	Extracellular domain in nidogen	13.91 3.49e-09 569-588
913	IPB003886D	Extracellular domain in nidogen	13.91 4.78e-13 696-715
913	IPB003886D	Extracellular domain in nidogen	13.91 9.41e-10 812-831
913	PR00010C	Type II EGF-like signature III	6.98 3.63e-09 574-584
913	PR00010C	Type II EGF-like signature III	6.98 8.02e-09 701-711
913	PR00010C	Type II EGF-like signature III	6.98 9.47e-13 817-827
913	PR00907B	Thrombomodulin signature II	11.50 9.43e-09 808-824
913	PR00907G	Thrombomodulin signature VII	10.43 8.09e-10 812-838
913	PR00907G	Thrombomodulin signature VII	10.43 8.85e-10 696-722
913	PR00907G	Thrombomodulin signature VII	10.43 9.27e-09 569-595
913	PR01303D	Plasmodium circumsporozoite protein signature IV	10.57 2.00e-09 187-204
914	IPB001398A	Macrophage migration inhibitory factor family	23.60 9.76e-26 2-41
914	IPB001398B	Macrophage migration inhibitory factor family	19.16 6.31e-18 42-68
914	IPB001398C	Macrophage migration inhibitory factor family	26.46 5.14e-12 70-116
915	IPB001398A	Macrophage migration inhibitory factor family	23.60 3.25e-20 2-41

Table 3B  
365

SEQ ID NO:	Database entry ID	Description	Results*
915	IPB001398B	Macrophage migration inhibitory factor family	19.16 2.06e-20 55-81
915	IPB001398C	Macrophage migration inhibitory factor family	26.46 7.39e-36 83-129
916	IPB000001B	Kringle	18.04 8.09e-19 286-310
916	IPB000001C	Kringle	20.41 2.23e-22 331-351
916	IPB000001D	Kringle	11.31 3.61e-17 469-485
916	IPB000001E	Kringle	16.02 8.58e-13 544-557
916	IPB000001F	Kringle	10.14 7.65e-11 567-577
916	IPB000001G	Kringle	29.29 2.80e-16 588-616
916	IPB000001H	Kringle	12.24 3.45e-13 628-638
916	IPB000177M	Apple domain	9.18 4.82e-10 545-579
916	IPB000177N	Apple domain	10.17 3.25e-20 618-652
916	IPB000177O	Apple domain	14.39 6.13e-17 653-681
916	IPB000561	EGF-like domain	4.89 4.21e-09 267-275
916	IPB000562	Type II fibronectin collagen-binding domain	10.39 9.05e-20 132-148
916	IPB001254A	Serine proteases, trypsin family	9.98 6.09e-14 469-485
916	IPB001254B	Serine proteases, trypsin family	15.01 3.50e-17 626-649
916	IPB001254C	Serine proteases, trypsin family	16.54 1.00e-18 656-675
916	IPB001774D	Delta serrate ligand	19.23 5.82e-09 152-198
916	IPB001881A	Calcium-binding EGF-like domain	8.72 1.00e-08 245-254
916	IPB001881A	Calcium-binding EGF-like domain	8.72 8.20e-09 164-173
916	PR00010A	Type II EGF-like signature I	12.91 6.54e-10 160-171
916	PR00011B	Type III EGF-like signature II	13.08 6.42e-09 179-197
916	PR00011B	Type III EGF-like signature II	13.08 9.38e-09 260-278
916	PR00013A	Fibronectin type II repeat signature I	12.70 4.60e-10 105-114
916	PR00013B	Fibronectin type II repeat signature II	11.78 5.25e-10 116-128
916	PR00013C	Fibronectin type II repeat signature III	12.27 9.47e-19 132-147
916	PR00018A	Kringle domain signature I	12.23 8.84e-12 286-301
916	PR00018B	Kringle domain signature II	16.62 1.75e-09 302-314
916	PR00018C	Kringle domain signature III	16.77 7.00e-21 331-351
916	PR00018D	Kringle domain signature IV	12.83 9.04e-10 356-367
916	PR00722A	Chymotrypsin serine protease family (S1) signature I	12.06 9.36e-14 470-485
916	PR00722C	Chymotrypsin serine protease family (S1) signature III	10.74 8.62e-14 625-637
916	PR01236A	Tumour necrosis factor beta (lymphotoxin-alpha) signature I	4.92 8.22e-09 22-38
917	IPB000034A	Laminin B	22.21 3.69e-10 419-454
917	IPB000034A	Laminin B	22.21 7.22e-13 1046-1081
917	IPB000034A	Laminin B	22.21 9.42e-10 291-326
917	IPB000034B	Laminin B	21.22 8.38e-11 808-846
917	IPB000034B	Laminin B	21.22 9.07e-19 443-481
917	IPB000034C	Laminin B	12.97 1.12e-10 542-560
917	IPB000034C	Laminin B	12.97 2.29e-09 805-823
917	IPB000034C	Laminin B	12.97 3.10e-10 492-510
917	IPB000034C	Laminin B	12.97 5.09e-09 897-915
917	IPB000034C	Laminin B	12.97 5.29e-12 1067-1085
917	IPB000034C	Laminin B	12.97 8.17e-14 955-973
917	IPB000034C	Laminin B	12.97 8.43e-11 853-871
917	IPB000034C	Laminin B	12.97 8.62e-12 1114-1132
917	IPB001134C	Netrin, C-terminus	17.82 9.77e-09 1014-1028
917	IPB001774C	Delta serrate ligand	18.25 9.27e-09 1105-1147

Table 3B  
366

SEQ ID NO:	Database entry ID	Description	Results*
917	IPB001886A	Laminin N-terminal (Domain VI)	8.51 9.44e-10 119-126
917	IPB001886B	Laminin N-terminal (Domain VI)	15.20 4.81e-24 269-291
917	IPB001886C	Laminin N-terminal (Domain VI)	24.54 2.14e-13 377-416
917	IPB001886C	Laminin N-terminal (Domain VI)	24.54 6.21e-11 1009-1048
917	IPB001886C	Laminin N-terminal (Domain VI)	24.54 6.55e-12 1067-1106
917	IPB001886C	Laminin N-terminal (Domain VI)	24.54 7.04e-11 955-994
917	IPB001886C	Laminin N-terminal (Domain VI)	24.54 8.03e-12 440-479
917	IPB001886C	Laminin N-terminal (Domain VI)	24.54 8.41e-32 312-351
917	IPB001886C	Laminin N-terminal (Domain VI)	24.54 8.81e-09 853-892
917	IPB001886D	Laminin N-terminal (Domain VI)	12.51 3.39e-13 893-911
917	IPB001886D	Laminin N-terminal (Domain VI)	12.51 5.86e-09 308-326
917	IPB001886D	Laminin N-terminal (Domain VI)	12.51 7.20e-14 1005-1023
917	IPB001886D	Laminin N-terminal (Domain VI)	12.51 9.64e-20 373-391
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 1.55e-11 438-454
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 2.20e-12 1065-1081
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 3.02e-11 490-506
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 6.00e-12 851-867
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 6.14e-15 540-556
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 6.33e-11 1112-1128
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 7.43e-11 310-326
917	IPB001886E	Laminin N-terminal (Domain VI)	10.90 9.07e-17 803-819
917	IPB003884C	Factor I membrane attack complex	13.00 3.35e-09 411-429
917	IPB003884C	Factor I membrane attack complex	13.00 3.87e-09 833-851
917	IPB003888L	FY-rich domain N-terminus	21.19 8.82e-10 339-385
917	PR00011A	Type III EGF-like signature I	14.05 2.02e-14 485-503
917	PR00011A	Type III EGF-like signature I	14.05 2.76e-17 1060-1078
917	PR00011A	Type III EGF-like signature I	14.05 3.35e-11 433-451
917	PR00011A	Type III EGF-like signature I	14.05 4.41e-13 798-816
917	PR00011A	Type III EGF-like signature I	14.05 4.72e-16 535-553
917	PR00011A	Type III EGF-like signature I	14.05 7.30e-15 846-864
917	PR00011A	Type III EGF-like signature I	14.05 7.50e-09 305-323
917	PR00011A	Type III EGF-like signature I	14.05 8.24e-16 1107-1125
917	PR00011B	Type III EGF-like signature II	13.08 4.17e-14 485-503
917	PR00011B	Type III EGF-like signature II	13.08 4.35e-15 1107-1125
917	PR00011B	Type III EGF-like signature II	13.08 4.86e-17 846-864
917	PR00011B	Type III EGF-like signature II	13.08 6.46e-12 798-816
917	PR00011B	Type III EGF-like signature II	13.08 6.96e-10 433-451
917	PR00011B	Type III EGF-like signature II	13.08 7.71e-15 535-553
917	PR00011B	Type III EGF-like signature II	13.08 8.00e-14 1060-1078
917	PR00011C	Type III EGF-like signature III	25.79 1.15e-09 455-483
917	PR00011C	Type III EGF-like signature III	25.79 6.21e-10 440-468
917	PR00011C	Type III EGF-like signature III	25.79 9.82e-12 817-845
917	PR00011D	Type III EGF-like signature IV	12.12 1.00e-16 1107-1125
917	PR00011D	Type III EGF-like signature IV	12.12 2.62e-14 846-864
917	PR00011D	Type III EGF-like signature IV	12.12 3.00e-13 535-553
917	PR00011D	Type III EGF-like signature IV	12.12 5.00e-13 485-503
917	PR00011D	Type III EGF-like signature IV	12.12 5.50e-14 1060-1078
917	PR00011D	Type III EGF-like signature IV	12.12 6.23e-11 433-451
917	PR00011D	Type III EGF-like signature IV	12.12 7.98e-12 798-816
918	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 1.00e-09 304-314
918	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 1.00e-10 132-142

Table 3B  
367

SEQ ID NO:	Database entry ID	Description	Results*
918	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 4.19e-10 344-354
918	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 5.74e-09 384-394
918	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 7.97e-10 91-101
918	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 3.12e-15 379-394
918	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 3.37e-14 299-314
918	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 3.45e-13 46-61
918	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 5.76e-15 339-354
918	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 7.55e-13 127-142
918	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 8.77e-13 86-101
918	IPB000561	EGF-like domain	4.89 5.50e-09 55-63
918	IPB000859	CUB domain	19.99 7.67e-17 721-756
918	IPB001881B	Calcium-binding EGF-like domain	12.28 1.47e-11 127-138
918	IPB001881B	Calcium-binding EGF-like domain	12.28 1.47e-11 299-310
918	IPB001881B	Calcium-binding EGF-like domain	12.28 5.00e-14 379-390
918	IPB001881B	Calcium-binding EGF-like domain	12.28 5.26e-11 86-97
918	IPB001881B	Calcium-binding EGF-like domain	12.28 7.00e-13 339-350
918	IPB001881B	Calcium-binding EGF-like domain	12.28 8.71e-10 46-57
918	IPB002557B	Chitin binding domain	12.64 3.08e-09 653-666
918	IPB003367A	Thrombospondin type 3 repeat	11.78 7.62e-14 54-74
918	IPB003886D	Extracellular domain in nidogen	13.91 1.54e-13 46-65
918	IPB003886D	Extracellular domain in nidogen	13.91 2.13e-11 339-358
918	IPB003886D	Extracellular domain in nidogen	13.91 9.17e-09 379-398
918	PR00010C	Type II EGF-like signature III	6.98 4.48e-10 132-142
918	PR00010C	Type II EGF-like signature III	6.98 5.61e-09 304-314
918	PR00010C	Type II EGF-like signature III	6.98 5.83e-09 384-394
918	PR00010C	Type II EGF-like signature III	6.98 6.05e-09 51-61
918	PR00010C	Type II EGF-like signature III	6.98 8.02e-09 91-101
918	PR00907B	Thrombomodulin signature II	11.50 9.16e-10 215-231
918	PR00907G	Thrombomodulin signature VII	10.43 6.84e-10 46-72
918	PR00907G	Thrombomodulin signature VII	10.43 7.51e-10 379-405
919	IPB001415A	Parathyroid hormone family	18.60 2.50e-37 80-114
919	IPB001415B	Parathyroid hormone family	26.47 1.00e-40 133-181
920	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 5.32e-11 144-154
920	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 6.76e-11 103-113
920	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 8.92e-11 17-27
920	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 3.05e-13 139-154
920	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 5.26e-14 12-27
920	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 9.59e-13 98-113
920	IPB000561	EGF-like domain	4.89 1.00e-09 252-260

Table 3B  
368

SEQ ID NO:	Database entry ID	Description	Results*
920	IPB000561	EGF-like domain	4.89 1.00e-09 67-75
920	IPB001881B	Calcium-binding EGF-like domain	12.28 1.43e-10 12-23
920	IPB001881B	Calcium-binding EGF-like domain	12.28 6.57e-10 139-150
920	IPB001881B	Calcium-binding EGF-like domain	12.28 7.86e-10 98-109
920	PR00010C	Type II EGF-like signature III	6.98 3.63e-09 103-113
920	PR00010C	Type II EGF-like signature III	6.98 7.59e-09 144-154
920	PR00011A	Type III EGF-like signature I	14.05 3.50e-09 331-349
920	PR00011A	Type III EGF-like signature I	14.05 6.09e-14 288-306
920	PR00011A	Type III EGF-like signature I	14.05 7.90e-12 376-394
920	PR00011B	Type III EGF-like signature II	13.08 1.00e-13 288-306
920	PR00011B	Type III EGF-like signature II	13.08 3.36e-12 376-394
920	PR00011B	Type III EGF-like signature II	13.08 3.59e-09 331-349
920	PR00011D	Type III EGF-like signature IV	12.12 2.50e-13 376-394
920	PR00011D	Type III EGF-like signature IV	12.12 2.86e-12 288-306
920	PR00011D	Type III EGF-like signature IV	12.12 4.05e-11 331-349
920	PR00764F	Complement C9 signature VI	15.74 5.10e-09 236-256
920	PR00907B	Thrombomodulin signature II	11.50 8.32e-10 54-70
921	IPB000827A	Small cytokines (intercrine/chemokine), C-C subfamily	8.72 5.24e-11 35-45
922	IPB000834A	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	18.14 1.77e-11 167-188
922	IPB000834B	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	13.51 5.50e-14 201-215
922	IPB000834C	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	17.20 6.66e-12 270-286
922	IPB000834D	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	18.95 3.25e-23 297-323
922	IPB000834E	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	9.80 3.33e-13 325-339
922	IPB000834F	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	12.40 9.18e-09 382-394
922	IPB000834G	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	14.46 6.00e-15 415-430
922	PR00765B	Carboxypeptidase A metalloprotease (M14) family signature II	14.48 8.04e-15 197-211
922	PR00765C	Carboxypeptidase A metalloprotease (M14) family signature III	10.88 5.50e-09 277-285
922	PR00765D	Carboxypeptidase A metalloprotease (M14) family signature IV	14.06 4.91e-14 330-343
923	IPB002181A	Fibrinogen beta and gamma chains C-terminal globular domain	18.44 7.00e-11 64-80
923	IPB002181B	Fibrinogen beta and gamma chains C-terminal globular domain	20.16 4.86e-32 93-129
923	IPB002181C	Fibrinogen beta and gamma chains C-terminal globular domain	15.87 6.88e-13 134-146
923	IPB002181D	Fibrinogen beta and gamma chains C-terminal globular domain	29.18 1.23e-18 153-193
923	IPB002181E	Fibrinogen beta and gamma chains C-terminal globular domain	27.75 9.71e-23 205-237
927	IPB000774B	Domain amino terminal to FKBP-type peptidyl-prolyl isomerase	11.42 1.29e-12 54-81
927	IPB000774C	Domain amino terminal to FKBP-type peptidyl-prolyl isomerase	10.28 7.46e-19 84-113
927	IPB001179	FKBP-type peptidyl-prolyl cis-trans	22.74 4.60e-26 80-113



Table 3B  
369

SEQ ID NO:	Database entry ID	Description	Results*
		isomerase (PPIase)	
932	IPB001359H	Synapsin	22.58 6.66e-10 176-226
936	PR00436A	Interleukin-8 signature I	15.20 8.23e-09 3-26
937	IPB000308B	14-3-3 proteins	12.24 2.33e-40 77-116
937	IPB000308C	14-3-3 proteins	18.31 1.00e-36 143-192
937	IPB000308D	14-3-3 proteins	12.57 1.69e-35 215-269
937	PR00305A	14-3-3 protein zeta signature I	9.70 9.18e-33 77-106
937	PR00305B	14-3-3 protein zeta signature II	11.03 4.52e-22 126-150
937	PR00305C	14-3-3 protein zeta signature III	9.89 5.50e-28 159-181
937	PR00305D	14-3-3 protein zeta signature IV	17.40 4.81e-27 194-220
937	PR00305F	14-3-3 protein zeta signature VI	15.40 1.00e-36 248-277
938	IPB002048	EF-hand family	7.91 5.24e-10 123-135
939	PR01353A	Glucagon family receptor signature I	8.87 4.86e-10 5-18
940	IPB001875B	Death effector domain	13.21 1.00e-08 500-509
940	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 9.65e-09 182-219
940	IPB003530A	Long hematopoietin receptor, soluble alpha chains	23.64 8.76e-09 339-393
941	IPB000483	Leucine rich repeat C-terminal domain	11.18 8.11e-14 242-256
941	PR00019B	Leucine-rich repeat signature II	11.42 8.91e-09 135-148
941	PR00364D	Disease resistance protein signature IV	14.89 4.60e-09 84-100
944	IPB000215A	Serpins	13.01 9.31e-15 200-223
944	IPB000215B	Serpins	9.87 4.79e-12 296-308
944	IPB000215C	Serpins	13.90 5.18e-11 323-337
944	IPB000215D	Serpins	15.35 9.27e-11 430-456
944	IPB000215E	Serpins	15.36 1.00e-15 508-532
944	IPB000895A	Transthyretin precursor	12.60 4.04e-09 39-58
944	PR00743B	Glycosyl hydrolase family 36 signature II	14.95 8.83e-09 240-260
945	IPB002905C	N2,N2-dimethylguanosine tRNA methyltransferase	19.39 5.00e-11 115-136
947	IPB000306	FYVE Zn-finger, rabphilin/VPS27/FAB1 type	8.96 3.25e-18 552-564
947	IPB001841	RING finger	10.69 3.37e-09 559-568
950	IPB001304A	C-type lectin domain	17.98 5.50e-11 117-141
950	PR00356D	Type II antifreeze protein signature IV	12.93 7.41e-09 174-190
951	IPB001304A	C-type lectin domain	17.98 2.29e-15 202-226
951	PR00356A	Type II antifreeze protein signature I	13.29 6.70e-10 197-209
951	PR00356B	Type II antifreeze protein signature II	12.71 4.63e-10 209-226
953	IPB001359H	Synapsin	22.58 4.43e-10 16-66
953	PR00910A	Luteovirus ORF6 protein signature I	2.74 3.46e-09 107-119
954	PR00901H	Pheromone B alpha-1 receptor signature VIII	14.75 4.05e-09 56-66
955	PR00436A	Interleukin-8 signature I	15.20 4.39e-09 36-59
956	IPB003406B	Core-2/I-Branching enzyme	11.63 9.31e-37 94-128
956	IPB003406C	Core-2/I-Branching enzyme	11.36 7.26e-37 145-184
956	IPB003406D	Core-2/I-Branching enzyme	23.29 1.82e-33 191-230
956	IPB003406E	Core-2/I-Branching enzyme	22.75 3.87e-29 251-292
958	IPB000130	Neutral zinc metalloproteases, zinc-binding region	5.86 6.54e-11 236-246
958	IPB001818A	Matrixin	14.60 7.35e-28 88-117
958	IPB001818B	Matrixin	26.48 7.75e-39 136-177
958	IPB001818C	Matrixin	24.38 1.00e-40 180-225
958	IPB001818D	Matrixin	14.91 5.34e-30 230-261
958	IPB001818E	Matrixin	8.86 4.67e-13 269-282

Table 3B  
370

SEQ ID NO:	Database entry ID	Description	Results*
958	IPB001818F	Matrixin	11.19 6.85e-22 305-325
958	IPB001818G	Matrixin	14.71 2.61e-12 342-354
958	IPB001818H	Matrixin	15.46 2.89e-28 408-434
958	IPB002870E	Reprolysin family propeptide	11.90 5.22e-10 238-250
958	PR00138A	Matrixin signature I	12.54 6.79e-16 108-121
958	PR00138B	Matrixin signature II	14.84 1.56e-17 155-170
958	PR00138C	Matrixin signature III	20.07 5.74e-28 178-206
958	PR00138D	Matrixin signature IV	14.57 7.86e-27 236-261
958	PR00138E	Matrixin signature V	7.10 4.96e-13 269-282
958	PR00480B	Astacin family signature II	14.35 4.32e-11 231-249
958	PR00997G	Fragilysin metalloproteinase (M10C) enterotoxin signature VII	11.64 8.53e-09 242-258
962	IPB000953	Chromo domain	14.73 2.29e-17 63-81
962	IPB003606E	N-terminal to some SET domains	9.57 2.89e-12 226-235
962	IPB003606G	N-terminal to some SET domains	13.63 4.00e-10 262-290
962	PR00504C	Chromodomain signature III	12.01 1.97e-10 72-84
965	IPB001729	Surfactant associated polypeptide SP-C palmitoylation sites	9.43 9.11e-09 19-52
967	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 9.18e-09 118-129
967	IPB000276D	Rhodopsin-like GPCR superfamily	9.40 9.22e-10 229-245
967	IPB001133A	NADH-ubiquinone oxidoreductase chain 4L	13.73 5.50e-10 89-119
967	PR00237A	Rhodopsin-like GPCR superfamily signature I	9.81 8.31e-09 26-50
967	PR00245B	Olfactory receptor signature II	13.73 3.57e-09 129-141
967	PR00245D	Olfactory receptor signature IV	9.34 2.59e-13 183-192
967	PR00245E	Olfactory receptor signature V	8.96 5.15e-16 230-241
967	PR00534A	Melanocortin receptor family signature I	12.77 2.83e-09 51-63
968	IPB002328A	Zinc-containing alcohol dehydrogenase	20.74 6.65e-09 179-210
968	IPB002328B	Zinc-containing alcohol dehydrogenase	14.71 9.03e-12 223-248
968	IPB002364	Quinone oxidoreductase/zeta-crystallin	20.78 6.67e-12 223-250
970	IPB001017A	Dehydrogenase, E1 component	11.25 7.35e-13 126-138
970	IPB001017B	Dehydrogenase, E1 component	25.91 2.64e-27 152-184
970	IPB001017C	Dehydrogenase, E1 component	15.29 7.00e-22 200-219
970	IPB001017D	Dehydrogenase, E1 component	18.59 1.45e-40 256-294
970	IPB001017E	Dehydrogenase, E1 component	10.22 3.65e-15 312-325
970	IPB001017F	Dehydrogenase, E1 component	17.40 9.31e-23 340-360
975	PR01281D	Corticotropin releasing factor type 2 receptor signature IV	9.32 8.52e-09 388-403
977	PR00457A	Animal haem peroxidase signature I	15.27 6.79e-12 162-173
977	PR00457B	Animal haem peroxidase signature II	12.43 1.26e-15 216-231
977	PR00457C	Animal haem peroxidase signature III	18.81 1.51e-14 366-384
977	PR00457D	Animal haem peroxidase signature IV	18.35 5.50e-22 384-404
977	PR00457E	Animal haem peroxidase signature V	19.97 4.91e-26 409-435
977	PR00457F	Animal haem peroxidase signature VI	14.42 6.05e-10 461-471
977	PR00457G	Animal haem peroxidase signature VII	14.17 3.25e-18 590-610
977	PR00457H	Animal haem peroxidase signature VIII	14.82 6.19e-13 684-698
978	IPB001271	Mammalian defensin	19.97 6.36e-13 95-123
985	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 8.63e-10 145-160
985	IPB002174A	Furin-like cysteine rich region	30.51 9.03e-10 26-57
987	PR00019B	Leucine-rich repeat signature II	11.42 8.33e-10 124-137
987	PR00019B	Leucine-rich repeat signature II	11.42 8.33e-10 316-329
988	IPB000392A	NifH/frxC family	9.08 2.98e-11 107-122
988	IPB000808A	Mrp family	23.51 8.67e-37 101-145

Table 3B  
371

SEQ ID NO:	Database entry ID	Description	Results*
988	IPB000808B	Mrp family	15.70 7.48e-36 191-227
988	IPB000808C	Mrp family	23.57 8.77e-29 250-285
988	IPB000808D	Mrp family	28.05 6.40e-24 297-337
988	IPB001984C	ATP-dependent serine proteases, Lon family	12.76 9.82e-09 103-139
988	IPB003348A	Anion-transporting ATPase	20.06 7.75e-15 106-143
988	PR00091A	Nitrogenase component II signature I	7.81 4.38e-09 109-122
988	PR00094A	Adenylate kinase signature I	9.62 9.86e-09 107-120
988	PR00364A	Disease resistance protein signature I	8.29 7.59e-10 105-120
989	IPB001007B	von Willebrand factor, type C repeat	10.03 4.79e-09 259-268
989	IPB001846B	von Willebrand factor type D domain	10.86 1.82e-12 489-503
989	IPB002919B	Trypsin Inhibitor-like cysteine rich domain	21.14 6.45e-10 321-343
989	PR00163B	Rubredoxin signature II	15.17 9.25e-09 215-231
992	IPB000353B	Class II histocompatibility antigen, beta chain, beta-1 domain	19.16 8.48e-10 133-182
992	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 1.38e-11 154-176
994	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 6.48e-11 86-123
994	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 8.04e-11 267-304
995	IPB001283A	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7	16.26 6.67e-10 29-48
995	IPB001283B	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7	13.09 2.33e-12 57-72
995	IPB001283E	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7	13.34 5.50e-20 127-147
995	IPB001283F	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7	10.49 6.48e-11 160-170
995	IPB003106E	Homeobox associated leucine zipper	13.73 6.90e-09 25-55
995	PR00837A	Allergen V5/Tpx-1 family signature I	14.69 1.56e-14 57-75
995	PR00837C	Allergen V5/Tpx-1 family signature III	16.31 5.50e-20 126-142
995	PR00837D	Allergen V5/Tpx-1 family signature IV	11.13 3.70e-11 160-173
995	PR00838D	Venom allergen 5 signature IV	8.59 3.57e-09 57-75
995	PR00838G	Venom allergen 5 signature VII	15.39 8.60e-18 125-144
1000	IPB000353B	Class II histocompatibility antigen, beta chain, beta-1 domain	19.16 3.25e-13 21-70
1000	IPB001003B	MHC Class II, alpha chain, alpha-1 domain	14.72 8.38e-11 29-72
1000	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 3.63e-11 42-64
1000	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 8.31e-17 78-115
1002	IPB000306	FYVE Zn-finger, rabphilin/VPS27/FAB1 type	8.96 9.74e-09 181-193
1002	IPB000822	Zinc finger, C2H2 type	14.67 1.60e-16 213-238
1002	IPB000822	Zinc finger, C2H2 type	14.67 2.38e-18 297-322
1002	IPB000822	Zinc finger, C2H2 type	14.67 2.50e-19 325-350
1002	IPB000822	Zinc finger, C2H2 type	14.67 7.23e-18 185-210
1002	IPB000822	Zinc finger, C2H2 type	14.67 7.23e-18 241-266
1002	IPB000822	Zinc finger, C2H2 type	14.67 7.75e-20 269-294
1002	IPB001275	DM DNA binding domain	19.17 1.85e-10 201-240
1002	IPB001275	DM DNA binding domain	19.17 4.65e-10 313-352
1002	IPB001275	DM DNA binding domain	19.17 5.50e-10 257-296
1002	IPB001275	DM DNA binding domain	19.17 6.24e-12 285-324

Table 3B  
372

SEQ ID NO:	Database entry ID	Description	Results*
1002	IPB001275	DM DNA binding domain	19.17 8.66e-10 229-268
1002	PR00048A	C2H2-type zinc finger signature I	9.94 1.00e-09 210-223
1002	PR00048A	C2H2-type zinc finger signature I	9.94 1.38e-09 182-195
1002	PR00048A	C2H2-type zinc finger signature I	9.94 1.41e-10 350-363
1002	PR00048A	C2H2-type zinc finger signature I	9.94 1.53e-12 238-251
1002	PR00048A	C2H2-type zinc finger signature I	9.94 1.95e-11 294-307
1002	PR00048A	C2H2-type zinc finger signature I	9.94 2.50e-09 322-335
1002	PR00048A	C2H2-type zinc finger signature I	9.94 6.79e-13 266-279
1002	PR00048B	C2H2-type zinc finger signature II	5.52 2.00e-09 198-207
1002	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-10 310-319
1002	PR00048B	C2H2-type zinc finger signature II	5.52 9.00e-09 338-347
1002	PR00048B	C2H2-type zinc finger signature II	5.52 9.36e-10 282-291
1003	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-17 467-492
1003	IPB000822	Zinc finger, C2H2 type	14.67 2.50e-20 523-548
1003	IPB000822	Zinc finger, C2H2 type	14.67 2.80e-16 262-287
1003	IPB000822	Zinc finger, C2H2 type	14.67 2.93e-17 439-464
1003	IPB000822	Zinc finger, C2H2 type	14.67 4.27e-21 383-408
1003	IPB000822	Zinc finger, C2H2 type	14.67 4.60e-16 663-688
1003	IPB000822	Zinc finger, C2H2 type	14.67 4.94e-15 635-660
1003	IPB000822	Zinc finger, C2H2 type	14.67 5.09e-21 411-436
1003	IPB000822	Zinc finger, C2H2 type	14.67 5.91e-21 495-520
1003	IPB000822	Zinc finger, C2H2 type	14.67 6.25e-19 327-352
1003	IPB000822	Zinc finger, C2H2 type	14.67 6.25e-20 579-604
1003	IPB000822	Zinc finger, C2H2 type	14.67 6.54e-18 355-380
1003	IPB000822	Zinc finger, C2H2 type	14.67 6.79e-17 299-324
1003	IPB000822	Zinc finger, C2H2 type	14.67 8.20e-22 607-632
1003	IPB000822	Zinc finger, C2H2 type	14.67 8.71e-17 551-576
1003	IPB001275	DM DNA binding domain	19.17 3.69e-12 371-410
1003	IPB001275	DM DNA binding domain	19.17 4.21e-11 399-438
1003	IPB001275	DM DNA binding domain	19.17 4.21e-11 483-522
1003	IPB001275	DM DNA binding domain	19.17 4.89e-10 539-578
1003	IPB001275	DM DNA binding domain	19.17 5.57e-13 595-634
1003	IPB001275	DM DNA binding domain	19.17 6.35e-10 455-494
1003	IPB001275	DM DNA binding domain	19.17 7.56e-11 511-550
1003	IPB001275	DM DNA binding domain	19.17 8.42e-10 343-382
1003	IPB001275	DM DNA binding domain	19.17 9.03e-10 651-690
1003	IPB001275	DM DNA binding domain	19.17 9.09e-09 287-326
1003	IPB001275	DM DNA binding domain	19.17 9.14e-13 567-606
1003	IPB001275	DM DNA binding domain	19.17 9.64e-10 315-354
1003	PR00048A	C2H2-type zinc finger signature I	9.94 1.00e-08 548-561
1003	PR00048A	C2H2-type zinc finger signature I	9.94 1.82e-10 259-272
1003	PR00048A	C2H2-type zinc finger signature I	9.94 2.06e-12 464-477
1003	PR00048A	C2H2-type zinc finger signature I	9.94 2.93e-13 604-617
1003	PR00048A	C2H2-type zinc finger signature I	9.94 3.25e-09 324-337
1003	PR00048A	C2H2-type zinc finger signature I	9.94 4.27e-10 520-533
1003	PR00048A	C2H2-type zinc finger signature I	9.94 4.86e-13 380-393
1003	PR00048A	C2H2-type zinc finger signature I	9.94 5.15e-14 576-589
1003	PR00048A	C2H2-type zinc finger signature I	9.94 6.54e-14 408-421
1003	PR00048A	C2H2-type zinc finger signature I	9.94 6.73e-10 436-449
1003	PR00048A	C2H2-type zinc finger signature I	9.94 7.23e-14 492-505
1003	PR00048A	C2H2-type zinc finger signature I	9.94 7.35e-12 352-365
1003	PR00048A	C2H2-type zinc finger signature I	9.94 8.88e-09 632-645
1003	PR00048A	C2H2-type zinc finger signature I	9.94 9.53e-11 660-673

Table 3B  
373

SEQ ID NO:	Database entry ID	Description	Results*
1003	PR00048B	C2H2-type zinc finger signature II	5.52 2.50e-09 676-685
1003	PR00048B	C2H2-type zinc finger signature II	5.52 3.57e-10 508-517
1003	PR00048B	C2H2-type zinc finger signature II	5.52 3.57e-10 536-545
1003	PR00048B	C2H2-type zinc finger signature II	5.52 3.57e-10 620-629
1003	PR00048B	C2H2-type zinc finger signature II	5.52 4.00e-09 592-601
1003	PR00048B	C2H2-type zinc finger signature II	5.52 4.86e-10 312-321
1003	PR00048B	C2H2-type zinc finger signature II	5.52 4.86e-10 396-405
1003	PR00048B	C2H2-type zinc finger signature II	5.52 6.14e-10 340-349
1003	PR00048B	C2H2-type zinc finger signature II	5.52 7.75e-11 424-433
1003	PR00048B	C2H2-type zinc finger signature II	5.52 8.00e-09 368-377
1003	PR00048B	C2H2-type zinc finger signature II	5.52 9.00e-09 564-573
1007	IPB001781	LIM domain	11.42 1.00e-09 536-546
1007	IPB001781	LIM domain	11.42 8.11e-12 469-479
1007	IPB001781	LIM domain	11.42 9.59e-11 411-421
1010	IPB000822	Zinc finger, C2H2 type	14.67 2.88e-09 156-181
1011	PR00456F	Ribosomal protein P2 signature VI	5.70 6.76e-12 76-87
1016	IPB002087A	Anti-proliferative protein	15.38 5.68e-13 14-28
1016	IPB002087B	Anti-proliferative protein	19.44 6.88e-20 122-155
1016	IPB002087C	Anti-proliferative protein	13.79 1.00e-23 180-200
1016	PR00310A	Anti-proliferative protein BTG1 family signature I	10.74 2.45e-28 16-40
1016	PR00310B	Anti-proliferative protein BTG1 family signature II	10.25 6.14e-24 123-152
1016	PR00310C	Anti-proliferative protein BTG1 family signature III	12.94 8.41e-36 153-182
1016	PR00310D	Anti-proliferative protein BTG1 family signature IV	9.99 4.71e-38 183-212
1016	PR00310E	Anti-proliferative protein BTG1 family signature V	14.23 2.06e-24 229-248
1017	IPB000306	FYVE Zn-finger, rabphilin/VPS27/FAB1 type	8.96 9.74e-09 491-503
1017	IPB000822	Zinc finger, C2H2 type	14.67 1.64e-17 411-436
1017	IPB000822	Zinc finger, C2H2 type	14.67 1.64e-17 523-548
1017	IPB000822	Zinc finger, C2H2 type	14.67 3.00e-24 383-408
1017	IPB000822	Zinc finger, C2H2 type	14.67 3.25e-20 355-380
1017	IPB000822	Zinc finger, C2H2 type	14.67 4.27e-21 467-492
1017	IPB000822	Zinc finger, C2H2 type	14.67 4.86e-11 215-240
1017	IPB000822	Zinc finger, C2H2 type	14.67 4.94e-14 271-296
1017	IPB000822	Zinc finger, C2H2 type	14.67 5.50e-09 121-146
1017	IPB000822	Zinc finger, C2H2 type	14.67 6.79e-17 439-464
1017	IPB000822	Zinc finger, C2H2 type	14.67 8.50e-20 327-352
1017	IPB000822	Zinc finger, C2H2 type	14.67 8.62e-18 495-520
1017	IPB000822	Zinc finger, C2H2 type	14.67 8.71e-10 149-174
1017	IPB001275	DM DNA binding domain	19.17 5.24e-11 343-382
1017	IPB001275	DM DNA binding domain	19.17 5.38e-10 483-522
1017	IPB001275	DM DNA binding domain	19.17 6.35e-09 315-354
1017	IPB001275	DM DNA binding domain	19.17 8.06e-09 427-466
1017	IPB001275	DM DNA binding domain	19.17 8.18e-09 455-494
1017	IPB001275	DM DNA binding domain	19.17 8.78e-10 511-550
1017	IPB001275	DM DNA binding domain	19.17 9.61e-11 371-410
1017	IPB001909	KRAB box	17.37 2.50e-33 44-78
1017	PR00048A	C2H2-type zinc finger signature I	9.94 1.00e-13 408-421
1017	PR00048A	C2H2-type zinc finger signature I	9.94 5.09e-10 324-337
1017	PR00048A	C2H2-type zinc finger signature I	9.94 5.24e-12 436-449

Table 3B  
374

SEQ ID NO:	Database entry ID	Description	Results*
1017	PR00048A	C2H2-type zinc finger signature I	9.94 6.54e-14 380-393
1017	PR00048A	C2H2-type zinc finger signature I	9.94 6.73e-10 492-505
1017	PR00048A	C2H2-type zinc finger signature I	9.94 6.79e-13 464-477
1017	PR00048A	C2H2-type zinc finger signature I	9.94 7.88e-12 352-365
1017	PR00048A	C2H2-type zinc finger signature I	9.94 8.13e-09 268-281
1017	PR00048A	C2H2-type zinc finger signature I	9.94 9.10e-15 520-533
1017	PR00048B	C2H2-type zinc finger signature II	5.52 2.50e-09 340-349
1017	PR00048B	C2H2-type zinc finger signature II	5.52 3.25e-11 396-405
1017	PR00048B	C2H2-type zinc finger signature II	5.52 3.25e-11 508-517
1017	PR00048B	C2H2-type zinc finger signature II	5.52 7.00e-11 368-377
1017	PR00048B	C2H2-type zinc finger signature II	5.52 9.36e-10 480-489
1020	IPB001781	LIM domain	11.42 1.69e-10 160-170
1020	IPB001781	LIM domain	11.42 1.95e-12 219-229
1020	IPB001781	LIM domain	11.42 5.26e-12 99-109
1022	PR01217G	Proline rich extensin signature VII	4.02 6.61e-09 130-155
1024	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-16 410-435
1024	IPB000822	Zinc finger, C2H2 type	14.67 3.25e-19 326-351
1024	IPB000822	Zinc finger, C2H2 type	14.67 3.25e-20 214-239
1024	IPB000822	Zinc finger, C2H2 type	14.67 3.77e-18 242-267
1024	IPB000822	Zinc finger, C2H2 type	14.67 4.00e-19 382-407
1024	IPB000822	Zinc finger, C2H2 type	14.67 4.75e-20 270-295
1024	IPB000822	Zinc finger, C2H2 type	14.67 4.94e-14 298-323
1024	IPB000822	Zinc finger, C2H2 type	14.67 5.50e-09 158-183
1024	IPB000822	Zinc finger, C2H2 type	14.67 6.00e-13 186-211
1024	IPB000822	Zinc finger, C2H2 type	14.67 8.07e-17 354-379
1024	IPB001275	DM DNA binding domain	19.17 1.68e-09 398-437
1024	IPB001275	DM DNA binding domain	19.17 3.07e-10 314-353
1024	IPB001275	DM DNA binding domain	19.17 5.37e-11 370-409
1024	IPB001275	DM DNA binding domain	19.17 5.74e-10 258-297
1024	IPB001275	DM DNA binding domain	19.17 6.10e-12 202-241
1024	IPB001275	DM DNA binding domain	19.17 8.12e-12 342-381
1024	IPB001275	DM DNA binding domain	19.17 9.66e-09 230-269
1024	IPB001502E	Endonuclease III	23.53 7.04e-09 252-279
1024	IPB001909	KRAB box	17.37 9.00e-33 17-51
1024	IPB002867C	Cysteine-rich domain (C6HC)	19.46 3.37e-09 323-340
1024	IPB002867C	Cysteine-rich domain (C6HC)	19.46 9.05e-09 267-284
1024	PR00048A	C2H2-type zinc finger signature I	9.94 1.82e-10 211-224
1024	PR00048A	C2H2-type zinc finger signature I	9.94 1.90e-15 351-364
1024	PR00048A	C2H2-type zinc finger signature I	9.94 2.23e-10 379-392
1024	PR00048A	C2H2-type zinc finger signature I	9.94 6.29e-12 323-336
1024	PR00048A	C2H2-type zinc finger signature I	9.94 7.43e-13 267-280
1024	PR00048A	C2H2-type zinc finger signature I	9.94 7.55e-10 295-308
1024	PR00048A	C2H2-type zinc finger signature I	9.94 8.58e-11 407-420
1024	PR00048A	C2H2-type zinc finger signature I	9.94 8.62e-14 239-252
1024	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-10 227-236
1024	PR00048B	C2H2-type zinc finger signature II	5.52 6.00e-09 423-432
1024	PR00048B	C2H2-type zinc finger signature II	5.52 8.07e-10 283-292
1024	PR00048B	C2H2-type zinc finger signature II	5.52 8.07e-10 339-348
1024	PR00048B	C2H2-type zinc finger signature II	5.52 8.07e-10 395-404
1025	PR00966H	Potyvirus nuclear inclusion A cysteine protease (C4) signature VIII	13.43 9.45e-09 215-234
1026	IPB000315A	B-box zinc finger superfamily	24.51 4.67e-12 19-42
1026	IPB001293A	TRAF-type zinc finger	26.83 8.34e-10 10-54

Table 3B  
375

SEQ ID NO:	Database entry ID	Description	Results*
1026	IPB003649A	B-Box C-terminal domain	13.56 7.92e-12 26-39
1026	PR00194D	Tropomyosin signature IV	9.54 7.21e-10 150-173
1026	PR01511F	Kv1.4 voltage-gated K <sup>+</sup> channel signature VI	5.19 4.54e-09 141-155
1029	IPB000006	Vertebrate metallothionein, family 1	13.41 6.13e-10 91-136
1029	IPB000006	Vertebrate metallothionein, family 1	13.41 6.23e-09 71-116
1029	IPB000118A	Granulin	8.25 8.84e-09 141-163
1029	IPB000359B	Cystine-knot domain	19.26 9.44e-09 181-199
1029	IPB000726B	Glycoside hydrolase family 19	5.39 2.54e-09 66-76
1029	IPB000967E	Zinc finger NF-X1 type	21.88 6.54e-09 105-145
1029	IPB000967E	Zinc finger NF-X1 type	21.88 7.83e-10 183-223
1029	IPB000967E	Zinc finger NF-X1 type	21.88 9.04e-09 136-176
1029	IPB001169K	Integrin beta, C-terminus	27.45 6.96e-09 107-149
1029	IPB001169K	Integrin beta, C-terminus	27.45 7.93e-09 56-98
1029	IPB001762A	Disintegrin	23.93 6.47e-09 185-225
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.84e-10 102-135
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.08e-13 164-197
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.25e-10 66-99
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.25e-10 92-125
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.46e-11 112-145
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.68e-11 169-202
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.89e-12 122-155
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.00e-14 159-192
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.39e-09 174-207
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.64e-11 206-239
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.19e-10 107-140
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.65e-09 91-124
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.04e-14 76-109
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.19e-13 81-114
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.71e-09 8-41
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.88e-15 117-150
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.00e-09 221-254
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.10e-09 67-100
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.48e-09 222-255
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.16e-09 216-249
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.33e-12 196-229
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.38e-13 71-104
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.84e-09 197-230
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.03e-09 86-119
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.28e-10 154-187
1029	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.32e-09 57-90
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.13e-09 70-84
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 2.01e-09 136-150
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 3.15e-10 22-36
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.94e-09 215-229
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.20e-09 105-119
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.58e-09 163-177
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 7.85e-10 210-224
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 8.23e-09 234-248
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 8.66e-10 95-109
1029	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.62e-09 121-135
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.10e-14 31-74
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.11e-17 69-112

Table 3B  
376

SEQ ID NO:	Database entry ID	Description	Results*
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.27e-11 27-70
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.27e-11 59-102
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.39e-13 199-242
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.64e-11 114-157
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.72e-15 104-147
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.79e-10 204-247
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.22e-12 100-143
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.35e-09 135-178
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.43e-09 177-220
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.80e-18 115-158
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.16e-19 162-205
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.20e-14 120-163
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.45e-13 94-137
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.79e-15 167-210
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.94e-09 85-128
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.03e-13 26-69
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.42e-13 64-107
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.72e-13 95-138
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.83e-15 84-127
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.00e-14 172-215
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.37e-09 183-226
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.38e-20 209-252
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.50e-12 219-262
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.71e-09 147-190
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.79e-09 70-113
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.09e-13 99-142
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.33e-10 90-133
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.42e-10 32-75
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.14e-09 142-185
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.16e-13 130-173
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.38e-19 157-200
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.39e-20 79-122
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.15e-09 16-59
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.15e-09 80-123
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.29e-19 110-153
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.36e-11 109-152
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.41e-12 125-168
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.45e-15 89-132
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.57e-09 189-232
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.86e-15 11-54
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.92e-13 152-195
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.13e-17 21-64
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.16e-21 74-117
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.27e-11 6-49
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.71e-13 36-79
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.75e-09 41-84
1029	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.80e-14 105-148
1029	IPB002919A	Trypsin Inhibitor-like cysteine rich domain	15.56 5.71e-09 24-36
1031	IPB002130A	Cyclophilin-type peptidyl-prolyl cis-trans isomerase	13.87 2.17e-15 18-39
1031	IPB002130B	Cyclophilin-type peptidyl-prolyl cis-trans isomerase	21.20 5.85e-39 48-86
1031	IPB002130C	Cyclophilin-type peptidyl-prolyl cis-trans	16.92 1.00e-40 98-135



Table 3B  
377

SEQ ID NO:	Database entry ID	Description	Results*
		isomerase	
1031	PR00153A	Cyclophilin peptidyl-prolyl cis-trans isomerase signature I	14.60 2.71e-10 24-39
1031	PR00153B	Cyclophilin peptidyl-prolyl cis-trans isomerase signature II	12.94 6.29e-17 53-65
1031	PR00153C	Cyclophilin peptidyl-prolyl cis-trans isomerase signature III	10.79 2.64e-20 96-111
1031	PR00153D	Cyclophilin peptidyl-prolyl cis-trans isomerase signature IV	10.73 1.47e-15 111-123
1031	PR00153E	Cyclophilin peptidyl-prolyl cis-trans isomerase signature V	8.39 1.43e-15 124-139
1032	PR00010C	Type II EGF-like signature III	6.98 5.17e-09 45-55
1035	IPB000822	Zinc finger, C2H2 type	14.67 1.69e-18 274-299
1035	IPB000822	Zinc finger, C2H2 type	14.67 2.29e-17 102-127
1035	IPB000822	Zinc finger, C2H2 type	14.67 2.29e-17 302-327
1035	IPB000822	Zinc finger, C2H2 type	14.67 3.77e-18 190-215
1035	IPB000822	Zinc finger, C2H2 type	14.67 3.81e-15 414-439
1035	IPB000822	Zinc finger, C2H2 type	14.67 4.75e-19 218-243
1035	IPB000822	Zinc finger, C2H2 type	14.67 4.75e-19 442-467
1035	IPB000822	Zinc finger, C2H2 type	14.67 5.50e-15 162-187
1035	IPB000822	Zinc finger, C2H2 type	14.67 6.25e-19 386-411
1035	IPB000822	Zinc finger, C2H2 type	14.67 6.25e-20 358-383
1035	IPB000822	Zinc finger, C2H2 type	14.67 6.50e-13 246-271
1035	IPB000822	Zinc finger, C2H2 type	14.67 6.79e-17 330-355
1035	IPB000822	Zinc finger, C2H2 type	14.67 7.00e-19 470-495
1035	IPB000822	Zinc finger, C2H2 type	14.67 7.23e-18 498-523
1035	IPB000822	Zinc finger, C2H2 type	14.67 8.07e-17 526-551
1035	IPB000822	Zinc finger, C2H2 type	14.67 8.50e-19 134-159
1035	IPB001275	DM DNA binding domain	19.17 1.34e-09 206-245
1035	IPB001275	DM DNA binding domain	19.17 2.70e-10 178-217
1035	IPB001275	DM DNA binding domain	19.17 3.55e-10 374-413
1035	IPB001275	DM DNA binding domain	19.17 4.08e-09 346-385
1035	IPB001275	DM DNA binding domain	19.17 5.86e-10 458-497
1035	IPB001275	DM DNA binding domain	19.17 8.78e-10 318-357
1035	IPB001275	DM DNA binding domain	19.17 9.15e-10 290-329
1035	IPB001275	DM DNA binding domain	19.17 9.88e-10 486-525
1035	IPB001909	KRAB box	17.37 3.86e-28 6-40
1035	PR00048A	C2H2-type zinc finger signature I	9.94 1.41e-10 187-200
1035	PR00048A	C2H2-type zinc finger signature I	9.94 2.29e-13 299-312
1035	PR00048A	C2H2-type zinc finger signature I	9.94 2.42e-11 131-144
1035	PR00048A	C2H2-type zinc finger signature I	9.94 2.64e-10 383-396
1035	PR00048A	C2H2-type zinc finger signature I	9.94 3.45e-10 411-424
1035	PR00048A	C2H2-type zinc finger signature I	9.94 4.18e-12 495-508
1035	PR00048A	C2H2-type zinc finger signature I	9.94 4.21e-13 523-536
1035	PR00048A	C2H2-type zinc finger signature I	9.94 5.09e-10 551-564
1035	PR00048A	C2H2-type zinc finger signature I	9.94 6.14e-13 327-340
1035	PR00048A	C2H2-type zinc finger signature I	9.94 6.14e-13 99-112
1035	PR00048A	C2H2-type zinc finger signature I	9.94 6.54e-14 215-228
1035	PR00048A	C2H2-type zinc finger signature I	9.94 6.54e-14 467-480
1035	PR00048A	C2H2-type zinc finger signature I	9.94 6.82e-12 243-256
1035	PR00048A	C2H2-type zinc finger signature I	9.94 7.92e-14 355-368
1035	PR00048A	C2H2-type zinc finger signature I	9.94 8.58e-11 159-172
1035	PR00048A	C2H2-type zinc finger signature I	9.94 9.05e-11 439-452
1035	PR00048A	C2H2-type zinc finger signature I	9.94 9.36e-13 271-284

Table 3B  
378

SEQ ID NO:	Database entry ID	Description	Results*
1035	PR00048B	C2H2-type zinc finger signature II	5.52 1.00e-08 483-492
1035	PR00048B	C2H2-type zinc finger signature II	5.52 5.00e-12 115-124
1035	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-10 203-212
1035	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-10 287-296
1035	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-11 455-464
1035	PR00048B	C2H2-type zinc finger signature II	5.52 7.50e-09 371-380
1035	PR00048B	C2H2-type zinc finger signature II	5.52 8.00e-09 147-156
1035	PR00048B	C2H2-type zinc finger signature II	5.52 8.50e-09 399-408
1038	IPB001902A	Sulfate transporter	10.29 9.57e-10 310-319
1038	IPB001902B	Sulfate transporter	21.91 3.43e-31 340-390
1038	IPB001902C	Sulfate transporter	23.62 6.44e-25 429-480
1038	IPB002645B	STAS domain	13.24 1.43e-11 179-193
1038	IPB002645B	STAS domain	13.24 1.43e-11 622-636
1039	IPB000626	Ubiquitin domain	21.25 7.65e-09 56-110
1041	IPB000729A	PMP-22/EMP/MP20 family	19.70 9.05e-09 47-68
1042	IPB003530A	Long hematopoietin receptor, soluble alpha chains	23.64 3.01e-09 362-416
1042	PR01474F	Vascular cell adhesion molecule-1 (VCAM-1) signature VI	14.81 9.48e-09 389-402
1043	IPB001526A	Ly-6/u-PAR domain	13.24 3.65e-13 73-88
1043	IPB001526C	Ly-6/u-PAR domain	13.04 6.82e-13 142-157
1043	IPB003605A	GS motif preceding kinase domain in TGF beta receptor	13.79 3.93e-09 149-160
1047	IPB000353B	Class II histocompatibility antigen, beta chain, beta-1 domain	19.16 5.99e-10 203-252
1047	IPB001039A	Major histocompatibility complex protein, Class I	17.17 1.00e-40 15-68
1047	IPB001039B	Major histocompatibility complex protein, Class I	27.55 1.00e-40 96-147
1047	IPB001039C	Major histocompatibility complex protein, Class I	19.82 1.00e-40 177-230
1047	IPB001039D	Major histocompatibility complex protein, Class I	16.49 4.43e-38 255-309
1047	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 3.84e-15 224-246
1047	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 6.25e-25 261-298
1047	IPB003363E	Glycoprotein GG/GX	13.35 2.94e-11 308-340
1051	IPB000729D	PMP-22/EMP/MP20 family	18.96 7.30e-10 173-200
1052	IPB001526C	Ly-6/u-PAR domain	13.04 1.77e-09 46-61
1054	PR00258B	Speract receptor signature II	7.94 9.50e-16 738-749
1054	PR00258C	Speract receptor signature III	9.05 2.32e-10 753-763
1054	PR00258D	Speract receptor signature IV	14.29 6.63e-09 785-799
1054	PR00258E	Speract receptor signature V	14.06 8.45e-14 808-820
1055	IPB001429A	ATP P2X receptor	17.17 1.00e-38 12-50
1055	IPB001429B	ATP P2X receptor	5.49 2.71e-12 62-71
1055	IPB001429C	ATP P2X receptor	19.96 5.68e-31 88-118
1055	IPB001429D	ATP P2X receptor	20.63 2.38e-19 143-164
1055	IPB001429E	ATP P2X receptor	15.32 1.43e-30 178-205
1055	IPB001429F	ATP P2X receptor	20.01 5.88e-30 248-277
1055	IPB001429G	ATP P2X receptor	9.93 2.64e-14 291-300
1055	IPB001429H	ATP P2X receptor	12.67 7.00e-34 309-335
1055	IPB001675B	Glycosyltransferase family 29	15.84 9.26e-09 339-361
1055	PR01307A	P2X purinoceptor family signature I	10.92 2.80e-12 82-90

Table 3B  
379

SEQ ID NO:	Database entry ID	Description	Results*
1055	PR01307B	P2X purinoceptor family signature II	10.30 1.00e-15 157-168
1055	PR01307C	P2X purinoceptor family signature III	10.17 3.77e-15 243-255
1055	PR01307D	P2X purinoceptor family signature IV	10.26 1.56e-15 290-300
1055	PR01307E	P2X purinoceptor family signature V	15.20 2.29e-19 310-324
1055	PR01311A	P2X4 purinoceptor signature I	7.73 8.32e-10 1-8
1055	PR01311C	P2X4 purinoceptor signature III	9.00 1.90e-12 30-38
1055	PR01311D	P2X4 purinoceptor signature IV	12.87 1.12e-10 49-55
1055	PR01311E	P2X4 purinoceptor signature V	6.43 1.00e-09 139-146
1055	PR01311F	P2X4 purinoceptor signature VI	9.15 1.15e-10 208-216
1055	PR01311G	P2X4 purinoceptor signature VII	11.74 8.00e-14 280-290
1055	PR01311H	P2X4 purinoceptor signature VIII	12.12 1.00e-12 303-312
1055	PR01311I	P2X4 purinoceptor signature IX	5.25 1.00e-11 359-366
1059	PR00854A	Prostaglandin D receptor signature I	15.24 9.08e-18 6-20
1059	PR00854B	Prostaglandin D receptor signature II	7.30 8.15e-21 41-58
1059	PR00854C	Prostaglandin D receptor signature III	12.92 1.64e-12 93-104
1059	PR00854D	Prostaglandin D receptor signature IV	9.41 2.50e-18 185-200
1059	PR00854E	Prostaglandin D receptor signature V	10.50 4.65e-26 236-259
1059	PR00854F	Prostaglandin D receptor signature VI	12.83 9.68e-11 321-332
1059	PR00854G	Prostaglandin D receptor signature VII	10.66 1.78e-18 341-357
1059	PR00854H	Prostaglandin D receptor signature VIII	14.71 6.20e-17 369-389
1059	PR00856E	Prostacyclin (prostanoid IP) receptor signature V	9.82 1.72e-09 178-194
1060	IPB001898A	Sodium:sulfate symporter family	13.06 5.30e-21 49-72
1060	IPB001898B	Sodium:sulfate symporter family	13.18 8.41e-22 88-114
1060	IPB001898C	Sodium:sulfate symporter family	9.97 1.00e-23 132-151
1060	IPB001898D	Sodium:sulfate symporter family	12.13 1.00e-21 208-232
1060	IPB001898E	Sodium:sulfate symporter family	13.97 6.76e-21 248-272
1060	IPB001898F	Sodium:sulfate symporter family	18.66 6.04e-20 309-335
1060	IPB001898G	Sodium:sulfate symporter family	16.36 8.36e-24 405-434
1060	IPB001898H	Sodium:sulfate symporter family	21.26 1.00e-40 486-534
1063	IPB001039A	Major histocompatibility complex protein, Class I	17.17 8.89e-11 19-72
1064	IPB001171A	Ergosterol biosynthesis ERG4/ERG24 enzymes	11.99 1.55e-10 11-24
1064	IPB001171B	Ergosterol biosynthesis ERG4/ERG24 enzymes	14.81 6.88e-15 166-178
1064	IPB001171C	Ergosterol biosynthesis ERG4/ERG24 enzymes	26.44 4.75e-33 181-217
1064	IPB001171D	Ergosterol biosynthesis ERG4/ERG24 enzymes	20.69 1.00e-40 241-282
1064	IPB001171G	Ergosterol biosynthesis ERG4/ERG24 enzymes	25.35 1.00e-40 298-350
1065	IPB003388E	Reticulon	14.24 6.29e-09 930-972
1066	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 2.04e-09 264-301
1067	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 5.20e-12 105-116
1067	PR00237C	Rhodopsin-like GPCR superfamily signature III	14.77 6.73e-11 91-113
1067	PR00237E	Rhodopsin-like GPCR superfamily signature V	13.03 4.77e-10 175-198
1067	PR00245A	Olfactory receptor signature I	10.98 3.65e-09 79-90
1067	PR00245B	Olfactory receptor signature II	13.73 4.60e-09 116-128
1067	PR00245D	Olfactory receptor signature IV	9.34 1.53e-13 212-221
1067	PR00245E	Olfactory receptor signature V	8.96 3.30e-10 259-270

Table 3B  
380

SEQ ID NO:	Database entry ID	Description	Results*
1067	PR00534A	Melanocortin receptor family signature I	12.77 8.43e-09 38-50
1069	IPB000095E	PAK-box /P21-Rho-binding	17.62 3.52e-11 410-455
1069	IPB000961C	Protein kinase C-terminal domain	15.48 9.39e-14 409-443
1069	IPB001245A	Tyrosine kinase catalytic domain	22.45 6.54e-15 402-442
1069	IPB001245B	Tyrosine kinase catalytic domain	21.68 6.18e-09 467-505
1069	IPB003527C	MAP kinase	14.70 7.37e-10 394-442
1069	IPB003605A	GS motif preceding kinase domain in TGF beta receptor	13.79 2.00e-13 92-103
1069	IPB003605C	GS motif preceding kinase domain in TGF beta receptor	14.92 1.00e-40 302-349
1069	IPB003605D	GS motif preceding kinase domain in TGF beta receptor	12.41 1.00e-40 398-439
1069	IPB003605E	GS motif preceding kinase domain in TGF beta receptor	21.14 1.64e-40 441-480
1069	IPB003605F	GS motif preceding kinase domain in TGF beta receptor	23.21 1.00e-40 504-557
1069	PR00109E	Tyrosine kinase catalytic domain signature V	12.99 5.58e-09 549-571
1070	IPB000729C	PMP-22/EMP/MP20 family	37.83 1.71e-09 80-132
1070	IPB000729D	PMP-22/EMP/MP20 family	18.96 9.33e-10 156-183
1071	IPB000353B	Class II histocompatibility antigen, beta chain, beta-1 domain	19.16 7.52e-14 201-250
1071	IPB001039A	Major histocompatibility complex protein, Class I	17.17 2.80e-26 14-67
1071	IPB001039B	Major histocompatibility complex protein, Class I	27.55 5.50e-23 90-141
1071	IPB001039C	Major histocompatibility complex protein, Class I	19.82 7.75e-16 176-229
1071	IPB001039C	Major histocompatibility complex protein, Class I	19.82 7.94e-23 175-228
1071	IPB001039D	Major histocompatibility complex protein, Class I	16.49 7.43e-29 254-308
1071	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 6.82e-17 222-244
1071	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 6.40e-19 260-297
1072	IPB002146	ATP synthase B/B' CF(0)	21.39 6.94e-09 188-226
1072	PR00308B	Type I antifreeze protein signature II	3.38 8.99e-09 3-14
1072	PR00308C	Type I antifreeze protein signature III	2.79 6.92e-09 3-12
1072	PR00308C	Type I antifreeze protein signature III	2.79 9.59e-10 4-13
1072	PR00308C	Type I antifreeze protein signature III	2.79 9.59e-10 5-14
1072	PR00698E	C.elegans Srg family integral membrane protein signature V	14.65 2.76e-09 111-136
1072	PR01511D	Kv1.4 voltage-gated K <sup>+</sup> channel signature IV	3.91 2.29e-09 4-14
1072	PR01511D	Kv1.4 voltage-gated K <sup>+</sup> channel signature IV	3.91 2.88e-09 5-15
1073	IPB002617C	Vesicular monoamine transporter	16.23 5.03e-10 198-250
1073	PR01035E	Tetracycline resistance protein signature V	10.87 1.00e-08 208-230
1073	PR01035F	Tetracycline resistance protein signature VI	12.88 1.00e-08 499-519
1074	IPB002617C	Vesicular monoamine transporter	16.23 5.03e-10 286-338
1074	PR01035E	Tetracycline resistance protein signature V	10.87 1.00e-08 296-318
1074	PR01035F	Tetracycline resistance protein signature VI	12.88 1.00e-08 587-607
1075	IPB000906A	ZU5 domain	22.49 3.72e-09 58-100

Table 3B  
381

SEQ ID NO:	Database entry ID	Description	Results*
1075	IPB000906G	ZU5 domain	25.85 7.95e-10 164-212
1075	PR00806C	Vinculin signature III	10.57 8.56e-09 350-367
1076	PR00237C	Rhodopsin-like GPCR superfamily signature III	14.77 3.25e-09 188-210
1076	PR00237E	Rhodopsin-like GPCR superfamily signature V	13.03 4.86e-09 276-299
1076	PR00237F	Rhodopsin-like GPCR superfamily signature VI	14.34 1.25e-10 322-346
1077	IPB003662A	General substrate transporters	18.97 2.89e-09 139-171
1079	IPB000101A	Gamma-glutamyltranspeptidase	29.14 3.45e-14 352-391
1079	PR01210A	Gamma-glutamyltranspeptidase signature I	17.15 1.50e-13 130-155
1080	IPB000684L	Eukaryotic RNA polymerase II heptapeptide repeat	3.49 8.35e-09 1357-1399
1080	IPB001237C	43 Kd postsynaptic protein	5.44 4.63e-11 28-80
1080	IPB001359H	Synapsin	22.58 8.69e-09 1371-1421
1081	IPB002688A	Sterol/diacylglycerol O-acyltransferase	19.24 8.65e-20 259-281
1081	IPB002688B	Sterol/diacylglycerol O-acyltransferase	22.16 2.50e-25 308-347
1081	IPB002688C	Sterol/diacylglycerol O-acyltransferase	22.98 3.13e-20 321-368
1081	IPB002688D	Sterol/diacylglycerol O-acyltransferase	24.36 3.57e-33 380-415
1081	IPB002688E	Sterol/diacylglycerol O-acyltransferase	13.67 2.50e-18 433-453
1082	PR01547A	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature I	10.65 6.40e-24 53-74
1082	PR01547B	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature II	17.77 2.38e-19 97-116
1082	PR01547C	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature III	9.50 5.94e-23 118-139
1082	PR01547D	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature IV	12.09 5.50e-29 144-169
1082	PR01547E	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature V	10.42 1.95e-26 171-195
1082	PR01547F	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature VI	11.64 5.95e-25 258-278
1082	PR01547G	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature VII	7.91 1.00e-36 298-325
1082	PR01547H	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature VIII	9.70 1.00e-30 333-357
1082	PR01547I	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature IX	15.37 7.94e-28 408-438
1082	PR01547J	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature X	14.06 3.15e-23 439-460
1082	PR01547K	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature XI	13.61 7.87e-23 462-481
1082	PR01547L	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature XII	10.35 6.29e-19 486-502
1082	PR01547M	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature XIII	10.06 3.44e-12 1178-1196
1083	IPB000095E	PAK-box /P21-Rho-binding	17.62 3.15e-12 367-412
1083	IPB000961C	Protein kinase C-terminal domain	15.48 4.66e-14 366-400
1083	IPB001245A	Tyrosine kinase catalytic domain	22.45 5.50e-14 359-399
1083	IPB001245B	Tyrosine kinase catalytic domain	21.68 6.18e-09 424-462
1083	IPB001526C	Ly-6/u-PAR domain	13.04 1.00e-08 105-120
1083	IPB003527C	MAP kinase	14.70 8.81e-10 351-399
1083	IPB003605A	GS motif preceding kinase domain in TGF beta receptor	13.79 1.32e-11 112-123

Table 3B  
382

SEQ ID NO:	Database entry ID	Description	Results*
1083	IPB003605B	GS motif preceding kinase domain in TGF beta receptor	10.56 5.00e-28 229-255
1083	IPB003605C	GS motif preceding kinase domain in TGF beta receptor	14.92 1.00e-40 259-306
1083	IPB003605D	GS motif preceding kinase domain in TGF beta receptor	12.41 1.00e-40 355-396
1083	IPB003605E	GS motif preceding kinase domain in TGF beta receptor	21.14 8.88e-39 398-437
1083	IPB003605F	GS motif preceding kinase domain in TGF beta receptor	23.21 6.81e-40 461-514
1083	PR00109E	Tyrosine kinase catalytic domain signature V	12.99 2.15e-09 535-557
1083	PR00653D	Activin type II receptor signature IV	12.01 6.70e-11 385-406
1084	IPB000095E	PAK-box /P21-Rho-binding	17.62 3.15e-12 429-474
1084	IPB000961C	Protein kinase C-terminal domain	15.48 4.66e-14 428-462
1084	IPB001245A	Tyrosine kinase catalytic domain	22.45 5.50e-14 421-461
1084	IPB001245B	Tyrosine kinase catalytic domain	21.68 6.18e-09 486-524
1084	IPB001526C	Ly-6/u-PAR domain	13.04 1.00e-08 105-120
1084	IPB003527C	MAP kinase	14.70 8.81e-10 413-461
1084	IPB003605A	GS motif preceding kinase domain in TGF beta receptor	13.79 1.32e-11 112-123
1084	IPB003605B	GS motif preceding kinase domain in TGF beta receptor	10.56 5.00e-28 291-317
1084	IPB003605C	GS motif preceding kinase domain in TGF beta receptor	14.92 1.00e-40 321-368
1084	IPB003605D	GS motif preceding kinase domain in TGF beta receptor	12.41 1.00e-40 417-458
1084	IPB003605E	GS motif preceding kinase domain in TGF beta receptor	21.14 8.88e-39 460-499
1084	IPB003605F	GS motif preceding kinase domain in TGF beta receptor	23.21 6.81e-40 523-576
1084	PR00109E	Tyrosine kinase catalytic domain signature V	12.99 2.15e-09 597-619
1084	PR00653D	Activin type II receptor signature IV	12.01 6.70e-11 447-468
1085	PR00541C	Muscarinic M4 receptor signature III	8.06 7.73e-09 486-506
1086	PR00541C	Muscarinic M4 receptor signature III	8.06 7.73e-09 755-775
1087	PR01430A	Protease activated receptor 2 precursor signature I	8.72 8.20e-09 123-134
1088	IPB000322A	Glycoside hydrolase family 31	19.86 8.13e-23 151-176
1088	IPB000322B	Glycoside hydrolase family 31	14.24 9.64e-10 186-201
1088	IPB000322C	Glycoside hydrolase family 31	22.80 3.57e-20 225-250
1088	IPB000322D	Glycoside hydrolase family 31	17.98 6.57e-27 326-352
1088	IPB000322E	Glycoside hydrolase family 31	25.60 9.10e-33 366-407
1090	IPB000135A	High mobility group proteins HMG1 and HMG2	11.69 7.66e-14 707-761
1090	IPB000135B	High mobility group proteins HMG1 and HMG2	13.24 5.68e-21 680-724
1090	PR00886A	High mobility group (HMG1/HMG2) protein signature I	10.57 4.46e-10 710-732
1090	PR00886C	High mobility group (HMG1/HMG2) protein signature III	11.20 6.28e-13 696-714
1090	PR00887A	Structure-specific recognition protein signature I	12.71 1.82e-21 343-359
1090	PR00887B	Structure-specific recognition protein	10.24 1.35e-20 365-381

Table 3B  
383

SEQ ID NO:	Database entry ID	Description	Results*
		signature II	
1090	PR00887C	Structure-specific recognition protein signature III	13.32 1.47e-21 388-404
1090	PR00887D	Structure-specific recognition protein signature IV	16.42 2.73e-15 453-466
1090	PR00887E	Structure-specific recognition protein signature V	10.75 3.40e-20 480-498
1090	PR00887F	Structure-specific recognition protein signature VI	14.12 6.46e-21 498-515
1090	PR00887G	Structure-specific recognition protein signature VII	13.65 7.00e-18 521-537
1090	PR00887H	Structure-specific recognition protein signature VIII	12.27 3.53e-20 537-555
1091	IPB000215A	Serpins	13.01 4.77e-13 341-364
1091	IPB000215B	Serpins	9.87 9.68e-09 449-461
1091	IPB000215C	Serpins	13.90 7.11e-11 476-490
1091	IPB000215D	Serpins	15.35 3.88e-16 578-604
1091	IPB000215E	Serpins	15.36 6.21e-16 659-683
1092	IPB000126A	Serine proteases, V8 family	11.75 2.88e-10 357-372
1092	IPB000126B	Serine proteases, V8 family	12.50 1.51e-11 471-487
1092	PR00834D	HtrA/DegQ protease family signature IV	11.75 6.23e-09 466-483
1092	PR00839B	V8 serine protease family signature II	11.20 8.12e-10 357-374
1097	IPB000454	Eubacterial and plasma membrane ATP synthase subunit C	27.64 2.50e-28 79-132
1097	PR00124A	ATP synthase C subunit signature I	8.69 4.46e-13 75-94
1097	PR00124B	ATP synthase C subunit signature II	14.47 8.14e-11 96-111
1097	PR00124C	ATP synthase C subunit signature III	12.82 6.63e-17 113-138
1098	IPB000573	Aconitase C-terminal domain	25.70 4.86e-34 947-996
1098	IPB001030A	Aconitase (aconitate hydratase)	37.52 1.17e-13 452-500
1098	IPB001030A	Aconitase (aconitate hydratase)	37.52 5.06e-23 392-440
1098	IPB001030B	Aconitase (aconitate hydratase)	18.77 4.32e-18 665-688
1098	IPB001030C	Aconitase (aconitate hydratase)	18.16 3.65e-12 710-724
1098	PR00415A	Aconitase family signature I	13.25 7.00e-10 323-336
1098	PR00415C	Aconitase family signature III	14.58 2.75e-13 376-389
1098	PR00415D	Aconitase family signature IV	14.53 2.93e-15 390-405
1098	PR00415E	Aconitase family signature V	9.37 1.53e-12 452-465
1098	PR00415F	Aconitase family signature VI	11.36 6.14e-12 466-479
1098	PR00415G	Aconitase family signature VII	14.02 2.35e-14 548-562
1098	PR00415H	Aconitase family signature VIII	13.87 4.79e-12 613-624
1098	PR00415I	Aconitase family signature IX	14.82 8.43e-15 675-688
1099	PR01088H	Na <sup>+</sup> /H <sup>+</sup> exchanger isoform 6 (NHE6) signature VIII	16.89 1.00e-08 92-109
1100	PR01088H	Na <sup>+</sup> /H <sup>+</sup> exchanger isoform 6 (NHE6) signature VIII	16.89 1.00e-08 188-205
1101	PR01088H	Na <sup>+</sup> /H <sup>+</sup> exchanger isoform 6 (NHE6) signature VIII	16.89 1.00e-08 142-159
1102	PR00259D	Transmembrane four family signature IV	14.90 2.43e-09 65-91
1104	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 8.56e-10 124-161
1104	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 9.65e-09 227-264
1106	IPB001983A	Translationally controlled tumor protein	14.49 1.43e-25 1-27
1106	IPB001983B	Translationally controlled tumor protein	15.99 4.43e-20 45-70
1106	IPB001983C	Translationally controlled tumor protein	16.91 1.41e-24 83-109

Table 3B  
384

SEQ ID NO:	Database entry ID	Description	Results*
1106	IPB001983D	Translationally controlled tumor protein	22.98 4.09e-28 123-163
1107	IPB000117D	Kappa casein	10.18 8.71e-09 513-547
1108	IPB001737A	Ribosomal RNA adenine dimethylase	27.11 1.86e-09 110-155
1109	IPB002048	EF-hand family	7.91 1.00e-09 42-54
1110	IPB000272	ATP1G1/PLM/MAT8 family	14.36 8.64e-23 99-134
1110	IPB003403A	Herpesvirus immediate early protein	21.25 6.59e-09 39-61
1112	PR00261B	Low density lipoprotein (LDL) receptor signature II	15.12 6.29e-09 127-148
1112	PR00261D	Low density lipoprotein (LDL) receptor signature IV	16.87 7.09e-09 127-148
1112	PR00261F	Low density lipoprotein (LDL) receptor signature VI	15.46 3.91e-09 127-148
1112	PR00764B	Complement C9 signature II	12.47 1.36e-11 122-142
1115	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 4.32e-10 216-227
1115	PR00245A	Olfactory receptor signature I	10.98 6.70e-10 190-201
1115	PR00245B	Olfactory receptor signature II	13.73 2.03e-09 227-239
1115	PR00245C	Olfactory receptor signature III	14.65 3.54e-10 274-290
1116	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 8.36e-09 5-16
1116	IPB000276D	Rhodopsin-like GPCR superfamily	9.40 6.19e-09 355-371
1116	PR00245A	Olfactory receptor signature I	10.98 3.16e-11 165-176
1116	PR00245B	Olfactory receptor signature II	13.73 1.77e-09 16-28
1116	PR00245B	Olfactory receptor signature II	13.73 9.07e-10 202-214
1116	PR00245C	Olfactory receptor signature III	14.65 2.96e-15 249-265
1116	PR00245D	Olfactory receptor signature IV	9.34 8.58e-12 309-318
1116	PR00245E	Olfactory receptor signature V	8.96 2.20e-15 356-367
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 1.92e-10 300-315
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 1.92e-10 565-580
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 2.83e-11 1213-1228
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 3.90e-09 143-158
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 4.77e-09 448-463
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 7.83e-11 1336-1351
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 8.02e-10 1093-1108
1119	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 8.92e-12 964-979
1119	IPB001928	Endothelin and related toxins	14.31 3.00e-09 949-972
1119	IPB002861B	Reeler domain	10.50 5.95e-11 970-998
1119	IPB002861C	Reeler domain	23.17 1.00e-12 957-1011
1119	IPB002861C	Reeler domain	23.17 3.45e-09 136-190
1120	IPB000079	Nucleosomal binding domain found in HMG14 and HMG17	8.09 1.64e-37 18-48
1120	PR00925A	Nonhistone chromosomal protein HMG17 family signature I	5.67 2.29e-19 18-32
1120	PR00925B	Nonhistone chromosomal protein HMG17 family signature II	3.75 3.18e-16 34-46
1120	PR00925C	Nonhistone chromosomal protein HMG17 family signature III	5.23 7.45e-09 47-57
1120	PR00925D	Nonhistone chromosomal protein HMG17	7.38 2.00e-14 66-76



Table 3B  
385

SEQ ID NO:	Database entry ID	Description	Results*
		family signature IV	
1121	PR00219C	Synaptobrevin signature III	6.91 1.56e-09 89-108
1124	IPB001304A	C-type lectin domain	17.98 8.88e-19 203-227
1125	IPB002896F	Herpesvirus glycoprotein D	21.08 9.22e-09 374-409
1125	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 5.15e-09 323-360
1126	PR01007I	Flagellar hook-length control protein signature IX	8.84 9.69e-09 467-490
1131	IPB000008D	C2 domain	14.83 3.05e-11 1008-1026
1131	IPB000738A	WHEP-TRS domain	22.67 8.02e-09 1068-1093
1131	PR00360A	C2 domain signature I	15.18 8.36e-09 682-694
1132	IPB001135C	NADH-ubiquinone oxidoreductase 49Kd chain	25.46 9.79e-09 72-122
1133	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 2.20e-12 181-192
1133	IPB000276D	Rhodopsin-like GPCR superfamily	9.40 6.09e-10 344-360
1133	PR00237G	Rhodopsin-like GPCR superfamily signature VII	19.23 7.37e-10 334-360
1133	PR00245A	Olfactory receptor signature I	10.98 8.50e-10 155-166
1133	PR00245B	Olfactory receptor signature II	13.73 8.20e-13 192-204
1133	PR00245C	Olfactory receptor signature III	14.65 4.10e-13 239-255
1133	PR00245D	Olfactory receptor signature IV	9.34 7.63e-12 298-307
1133	PR00245E	Olfactory receptor signature V	8.96 4.77e-12 345-356
1134	IPB000425A	MIP family	9.37 7.33e-11 92-107
1134	IPB000425B	MIP family	22.94 1.75e-31 123-175
1134	IPB000425C	MIP family	10.67 7.28e-10 240-254
1134	IPB000425D	MIP family	8.10 9.67e-15 307-325
1134	PR00783A	Major intrinsic protein family signature I	11.82 5.50e-14 88-107
1134	PR00783B	Major intrinsic protein family signature II	15.34 1.67e-19 127-151
1134	PR00783C	Major intrinsic protein family signature III	13.32 2.76e-10 164-183
1134	PR00783E	Major intrinsic protein family signature V	16.25 8.11e-09 128-150
1134	PR00783E	Major intrinsic protein family signature V	16.25 8.13e-11 226-248
1134	PR00783F	Major intrinsic protein family signature VI	12.30 1.53e-14 308-328
1136	IPB002897A	Monocarboxylate transporter	16.12 4.90e-23 18-47
1136	IPB002897B	Monocarboxylate transporter	24.20 4.30e-32 72-126
1136	IPB002897C	Monocarboxylate transporter	13.63 4.46e-13 143-168
1136	IPB002897C	Monocarboxylate transporter	13.63 7.38e-12 177-202
1136	IPB002897D	Monocarboxylate transporter	30.44 5.85e-27 296-348
1136	IPB002897E	Monocarboxylate transporter	32.06 9.59e-24 407-458
1137	IPB000454	Eubacterial and plasma membrane ATP synthase subunit C	27.64 9.85e-11 94-147
1137	PR00122A	Vacuolar ATP synthase 16kDa subunit signature I	10.78 6.00e-18 30-54
1137	PR00122B	Vacuolar ATP synthase 16kDa subunit signature II	10.11 4.32e-26 56-80
1137	PR00122C	Vacuolar ATP synthase 16kDa subunit signature III	10.28 9.40e-32 104-130
1137	PR00122D	Vacuolar ATP synthase 16kDa subunit signature IV	10.03 7.00e-28 131-154
1137	PR00124C	ATP synthase C subunit signature III	12.82 7.58e-11 128-153
1138	IPB002220A	Dihydrodipicolinate synthetase	16.79 9.29e-09 41-53
1138	IPB002220B	Dihydrodipicolinate synthetase	14.56 6.84e-14 68-89
1138	IPB002220C	Dihydrodipicolinate synthetase	25.64 1.23e-24 108-155
1138	IPB002220D	Dihydrodipicolinate synthetase	10.31 6.56e-10 164-175
1138	IPB002220F	Dihydrodipicolinate synthetase	12.81 6.14e-10 234-247

Table 3B  
386

SEQ ID NO:	Database entry ID	Description	Results*
1138	PR00146D	Dihydrodipicolinate synthase signature IV	14.98 9.79e-11 163-180
1139	IPB001734A	Sodium:solute symporter family	14.69 5.76e-16 51-67
1139	IPB001734B	Sodium:solute symporter family	18.26 6.63e-27 166-206
1139	IPB001734C	Sodium:solute symporter family	18.21 7.75e-23 924-953
1140	IPB000074C	Apolipoprotein A1/A4/E	22.23 1.58e-09 385-422
1140	IPB002607B	Hydratase/decarboxylase	29.50 3.64e-10 242-290
1140	PR00122B	Vacuolar ATP synthase 16kDa subunit signature II	10.11 7.56e-09 493-517
1142	IPB000897A	GTP-binding signal recognition particle (SRP54) domain	9.15 9.87e-09 611-630
1143	IPB000219A	Dbl domain (dbl/cdc24 rhoGEF family)	10.71 3.00e-09 156-165
1143	IPB000276C	Rhodopsin-like GPCR superfamily	8.03 1.41e-10 278-289
1143	IPB000276D	Rhodopsin-like GPCR superfamily	9.40 2.17e-10 324-340
1143	PR00237A	Rhodopsin-like GPCR superfamily signature I	9.81 8.92e-11 63-87
1143	PR00237B	Rhodopsin-like GPCR superfamily signature II	12.45 4.08e-09 96-117
1143	PR00237C	Rhodopsin-like GPCR superfamily signature III	14.77 5.50e-12 140-162
1143	PR00237E	Rhodopsin-like GPCR superfamily signature V	13.03 8.46e-09 226-249
1143	PR00237F	Rhodopsin-like GPCR superfamily signature VI	14.34 9.33e-13 268-292
1143	PR00237G	Rhodopsin-like GPCR superfamily signature VII	19.23 6.04e-15 314-340
1143	PR01157A	P2 purinoceptor signature I	9.05 3.52e-11 110-122
1144	PR00049D	Wilm's tumour protein signature IV	0.00 2.07e-09 7-21
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 1.10e-11 266-290
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 1.92e-11 265-289
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 2.23e-11 270-294
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 3.00e-09 262-286
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 3.25e-10 271-295
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 4.55e-09 272-296
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 5.09e-09 260-284
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 7.07e-10 267-291
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 7.09e-09 259-283
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 7.40e-12 268-292
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 8.63e-10 263-287
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 8.98e-11 269-293
1145	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 9.24e-12 264-288
1145	IPB000637B	HMG-I and HMG-Y DNA-binding domain	14.21 8.64e-09 272-290

Table 3B  
387

SEQ ID NO:	Database entry ID	Description	Results*
		(A+T-hook)	
1145	IPB001202B	WW / rsp5 / WWP domain	12.18 6.40e-15 387-401
1145	IPB001422C	Neuromodulin (GAP-43)	16.82 8.65e-09 258-293
1145	IPB001580F	Calreticulin family	2.93 8.65e-09 277-286
1145	PR00403A	WW domain signature I	19.41 5.03e-10 373-386
1145	PR00403B	WW domain signature II	13.50 5.33e-15 387-401
1148	PR00360B	C2 domain signature II	11.64 1.29e-11 279-292
1150	IPB000685G	Ribulose biphosphate carboxylase, large chain	10.38 5.23e-09 45-99
1150	PR01237A	Tumour necrosis factor c (lymphotoxin-beta) signature I	6.39 8.27e-09 38-61
1152	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 7.63e-10 238-249
1152	PR00245A	Olfactory receptor signature I	10.98 1.30e-10 212-223
1152	PR00245B	Olfactory receptor signature II	13.73 5.50e-11 249-261
1152	PR00245C	Olfactory receptor signature III	14.65 6.87e-15 296-312
1152	PR00245E	Olfactory receptor signature V	8.96 2.04e-13 403-414
1152	PR00534A	Melanocortin receptor family signature I	12.77 7.52e-09 171-183
1153	IPB001680	G-protein beta WD-40 repeats	10.43 1.00e-11 456-467
1153	PR00320B	G protein beta WD-40 repeat signature II	12.82 7.55e-11 454-468
1153	PR00320C	G protein beta WD-40 repeat signature III	12.32 4.43e-11 454-468
1153	PR00962A	Lethal(2) giant larvae protein signature I	12.91 3.57e-22 17-35
1153	PR00962B	Lethal(2) giant larvae protein signature II	12.01 3.52e-29 310-332
1153	PR00962C	Lethal(2) giant larvae protein signature III	8.03 6.40e-23 362-382
1153	PR00962D	Lethal(2) giant larvae protein signature IV	9.91 4.16e-28 451-474
1153	PR00962E	Lethal(2) giant larvae protein signature V	8.81 3.67e-21 531-549
1153	PR00962F	Lethal(2) giant larvae protein signature VI	13.40 8.50e-24 568-587
1153	PR00962G	Lethal(2) giant larvae protein signature VII	14.51 6.54e-29 609-633
1153	PR00962H	Lethal(2) giant larvae protein signature VIII	13.53 3.53e-23 639-658
1153	PR00962I	Lethal(2) giant larvae protein signature IX	12.48 5.13e-23 708-727
1155	IPB002027A	Amino acid permease	18.88 8.82e-11 185-213
1155	IPB002027C	Amino acid permease	19.67 8.31e-21 385-423
1155	IPB002027D	Amino acid permease	22.00 6.79e-21 466-505
1157	PR01436C	NADH-ubiquinone oxidoreductase chain 2 signature III	9.53 7.75e-09 82-100
1158	IPB001428B	dUTPase	18.14 1.00e-14 107-125
1158	IPB001969A	Eukaryotic and viral aspartic protease active site	16.37 5.71e-09 195-211
1158	PR00783C	Major intrinsic protein family signature III	13.32 5.26e-17 31-50
1160	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 9.42e-10 1750-1760
1160	IPB000034A	Laminin B	22.21 1.85e-19 487-522
1160	IPB000034A	Laminin B	22.21 2.99e-10 1356-1391
1160	IPB000034A	Laminin B	22.21 3.84e-17 875-910
1160	IPB000034A	Laminin B	22.21 8.13e-10 1733-1768
1160	IPB000034A	Laminin B	22.21 8.56e-09 1263-1298
1160	IPB000034A	Laminin B	22.21 9.44e-09 533-568
1160	IPB000034B	Laminin B	21.22 1.23e-14 654-692
1160	IPB000034B	Laminin B	21.22 3.25e-31 1426-1464
1160	IPB000034B	Laminin B	21.22 3.33e-12 1287-1325
1160	IPB000034B	Laminin B	21.22 4.77e-10 1757-1795
1160	IPB000034B	Laminin B	21.22 4.99e-18 601-639
1160	IPB000034B	Laminin B	21.22 5.45e-09 511-549
1160	IPB000034C	Laminin B	12.97 3.09e-11 508-526
1160	IPB000034C	Laminin B	12.97 3.45e-09 319-337

Table 3B  
388

SEQ ID NO:	Database entry ID	Description	Results*
1160	IPB000034C	Laminin B	12.97 4.00e-11 443-461
1160	IPB000034C	Laminin B	12.97 4.05e-12 1377-1395
1160	IPB000034C	Laminin B	12.97 6.52e-13 554-572
1160	IPB000034C	Laminin B	12.97 6.56e-20 1754-1772
1160	IPB000034C	Laminin B	12.97 6.73e-09 1284-1302
1160	IPB000034C	Laminin B	12.97 8.01e-09 598-616
1160	IPB000034C	Laminin B	12.97 9.09e-11 651-669
1160	IPB000034C	Laminin B	12.97 9.31e-16 1653-1671
1160	IPB000561	EGF-like domain	4.89 5.50e-09 1754-1762
1160	IPB001001D	DNA polymerase III, beta chain	14.35 8.29e-09 2372-2395
1160	IPB001169K	Integrin beta, C-terminus	27.45 8.93e-11 1691-1733
1160	IPB001271	Mammalian defensin	19.97 5.91e-09 1728-1756
1160	IPB001774C	Delta serrate ligand	18.25 8.16e-09 545-587
1160	IPB001886A	Laminin N-terminal (Domain VI)	8.51 8.27e-09 124-131
1160	IPB001886B	Laminin N-terminal (Domain VI)	15.20 5.11e-21 286-308
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 1.00e-09 1284-1323
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 1.00e-11 1703-1742
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 1.23e-10 508-547
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 1.53e-14 1754-1793
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 2.97e-33 319-358
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 3.28e-13 391-430
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 4.19e-10 554-593
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 5.67e-10 1423-1462
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 6.33e-11 598-637
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 7.49e-10 1377-1416
1160	IPB001886C	Laminin N-terminal (Domain VI)	24.54 8.11e-11 651-690
1160	IPB001886D	Laminin N-terminal (Domain VI)	12.51 1.71e-09 439-457
1160	IPB001886D	Laminin N-terminal (Domain VI)	12.51 3.65e-17 1699-1717
1160	IPB001886D	Laminin N-terminal (Domain VI)	12.51 8.95e-10 1750-1768
1160	IPB001886D	Laminin N-terminal (Domain VI)	12.51 9.05e-22 387-405
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 1.21e-13 1282-1298
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 1.47e-09 317-333
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 4.82e-16 1375-1391
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 5.97e-09 649-665
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 6.74e-09 441-457
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 7.43e-13 1752-1768
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 7.61e-11 596-612
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 8.35e-11 506-522
1160	IPB001886E	Laminin N-terminal (Domain VI)	10.90 9.50e-20 552-568
1160	IPB003884C	Factor I membrane attack complex	13.00 3.22e-09 1403-1421
1160	PR00011A	Type III EGF-like signature I	14.05 1.68e-14 1416-1434
1160	PR00011A	Type III EGF-like signature I	14.05 1.69e-11 644-662
1160	PR00011A	Type III EGF-like signature I	14.05 3.88e-15 1277-1295
1160	PR00011A	Type III EGF-like signature I	14.05 3.93e-16 1370-1388
1160	PR00011A	Type III EGF-like signature I	14.05 3.95e-13 436-454
1160	PR00011A	Type III EGF-like signature I	14.05 5.00e-09 591-609
1160	PR00011A	Type III EGF-like signature I	14.05 5.05e-12 1318-1336
1160	PR00011A	Type III EGF-like signature I	14.05 5.43e-11 695-713
1160	PR00011A	Type III EGF-like signature I	14.05 6.82e-10 501-519
1160	PR00011A	Type III EGF-like signature I	14.05 8.83e-16 547-565
1160	PR00011B	Type III EGF-like signature II	13.08 1.00e-09 591-609
1160	PR00011B	Type III EGF-like signature II	13.08 1.66e-10 695-713
1160	PR00011B	Type III EGF-like signature II	13.08 4.18e-11 644-662

Table 3B  
389

SEQ ID NO:	Database entry ID	Description	Results*
1160	PR00011B	Type III EGF-like signature II	13.08 4.35e-15 1370-1388
1160	PR00011B	Type III EGF-like signature II	13.08 5.57e-12 436-454
1160	PR00011B	Type III EGF-like signature II	13.08 5.71e-17 547-565
1160	PR00011B	Type III EGF-like signature II	13.08 6.46e-12 1416-1434
1160	PR00011B	Type III EGF-like signature II	13.08 7.52e-13 1277-1295
1160	PR00011B	Type III EGF-like signature II	13.08 8.48e-11 1318-1336
1160	PR00011C	Type III EGF-like signature III	25.79 2.62e-09 651-679
1160	PR00011C	Type III EGF-like signature III	25.79 6.53e-10 1423-1451
1160	PR00011D	Type III EGF-like signature IV	12.12 1.18e-14 1370-1388
1160	PR00011D	Type III EGF-like signature IV	12.12 3.63e-10 695-713
1160	PR00011D	Type III EGF-like signature IV	12.12 4.96e-14 501-519
1160	PR00011D	Type III EGF-like signature IV	12.12 6.01e-09 591-609
1160	PR00011D	Type III EGF-like signature IV	12.12 7.65e-10 644-662
1160	PR00011D	Type III EGF-like signature IV	12.12 8.25e-17 547-565
1160	PR00011D	Type III EGF-like signature IV	12.12 8.33e-13 436-454
1160	PR00011D	Type III EGF-like signature IV	12.12 8.92e-14 1416-1434
1160	PR00011D	Type III EGF-like signature IV	12.12 9.17e-13 1277-1295
1160	PR00011D	Type III EGF-like signature IV	12.12 9.23e-09 1747-1765
1160	PR00877D	Plant PEC family metallothionein signature IV	6.45 1.89e-09 1308-1315
1161	IPB000034C	Laminin B	12.97 6.42e-10 205-223
1161	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 1.00e-10 158-173
1161	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 9.29e-09 196-211
1161	IPB000359A	Cystine-knot domain	23.24 4.86e-10 129-153
1161	IPB000561	EGF-like domain	4.89 9.25e-10 46-54
1161	IPB001762A	Disintegrin	23.93 9.65e-09 95-135
1161	IPB001774C	Delta serrate ligand	18.25 1.71e-31 37-79
1161	IPB001774C	Delta serrate ligand	18.25 9.17e-12 68-110
1161	IPB001862F	Membrane attack complex components/perforin/complement C9	29.39 2.65e-09 170-217
1161	IPB001862F	Membrane attack complex components/perforin/complement C9	29.39 8.53e-09 132-179
1161	IPB001886E	Laminin N-terminal (Domain VI)	10.90 9.67e-10 44-60
1161	PR00010A	Type II EGF-like signature I	12.91 1.27e-09 143-154
1161	PR00010C	Type II EGF-like signature III	6.98 2.54e-09 163-173
1161	PR00011A	Type III EGF-like signature I	14.05 6.88e-09 39-57
1161	PR00011B	Type III EGF-like signature II	13.08 7.88e-10 39-57
1161	PR00011D	Type III EGF-like signature IV	12.12 4.57e-12 39-57
1161	PR01217G	Proline rich extensin signature VII	4.02 7.79e-09 221-246
1162	IPB000471A	Interferon alpha, beta and delta family	27.36 2.86e-34 45-98
1162	PR00266A	Interferon alpha and beta subunit signature I	13.41 9.59e-14 67-79
1163	IPB000975D	Interleukin-1	24.45 5.55e-09 59-98
1163	IPB000975E	Interleukin-1	28.12 9.80e-09 103-142
1163	PR00264A	Interleukin-1 precursor family signature I	18.63 1.00e-08 62-82
1163	PR01360C	Interleukin-1 receptor antagonist precursor IL-1RA signature III	10.33 4.84e-11 65-82
1163	PR01360F	Interleukin-1 receptor antagonist precursor IL-1RA signature VI	14.44 3.11e-12 124-142
1164	IPB000072	Platelet-derived growth factor (PDGF)	13.51 4.00e-26 75-97
1164	IPB003498F	Probable DNA packing protein, C-terminus	23.31 7.55e-09 154-193
1165	IPB000975E	Interleukin-1	28.12 3.57e-12 95-134
1165	PR00264A	Interleukin-1 precursor family signature I	18.63 9.85e-09 55-75

Table 3B  
390

SEQ ID NO:	Database entry ID	Description	Results*
1165	PR00264C	Interleukin-1 precursor family signature III	19.37 4.90e-16 108-123
1165	PR01357F	Interleukin-1 alpha/beta precursor family signature VI	17.87 7.15e-10 108-123
1165	PR01360E	Interleukin-1 receptor antagonist precursor IL-1RA signature V	9.69 9.33e-13 95-115
1165	PR01360F	Interleukin-1 receptor antagonist precursor IL-1RA signature VI	14.44 9.86e-18 116-134
1166	IPB000276C	Rhodopsin-like GPCR superfamily	8.03 6.50e-11 33-44
1166	PR00237F	Rhodopsin-like GPCR superfamily signature VI	14.34 5.05e-16 23-47
1166	PR00237G	Rhodopsin-like GPCR superfamily signature VII	19.23 8.80e-13 70-96
1166	PR00427E	Interleukin-8 receptor signature V	7.33 3.84e-17 6-21
1166	PR00427F	Interleukin-8 receptor signature VI	12.53 2.63e-12 48-64
1166	PR00427G	Interleukin-8 receptor signature VII	10.47 8.53e-17 98-114
1166	PR00573D	Interleukin 8B receptor signature IV	15.06 1.85e-18 114-129
1167	PR00414A	Palmitoyl protein thioesterase signature I	10.61 3.44e-31 26-48
1167	PR00414B	Palmitoyl protein thioesterase signature II	11.87 6.79e-25 66-85
1167	PR00414C	Palmitoyl protein thioesterase signature III	13.39 1.90e-23 106-122
1167	PR00414D	Palmitoyl protein thioesterase signature IV	17.22 2.13e-27 151-170
1167	PR00414E	Palmitoyl protein thioesterase signature V	17.10 1.43e-27 185-204
1167	PR00414G	Palmitoyl protein thioesterase signature VII	12.98 2.36e-24 258-276
1168	IPB000215A	Serpins	13.01 4.18e-20 73-96
1168	IPB000215B	Serpins	9.87 9.68e-09 181-193
1168	IPB000215C	Serpins	13.90 7.11e-11 208-222
1168	IPB000215D	Serpins	15.35 3.88e-16 310-336
1168	IPB000215E	Serpins	15.36 1.56e-18 391-415
1170	IPB001271	Mammalian defensin	19.97 3.42e-21 103-131
1172	IPB000387	Tyrosine specific protein phosphatase and dual specificity protein phosphatase family	10.77 5.85e-13 439-449
1172	PR00700A	Protein tyrosine phosphatase signature I	6.05 1.00e-10 295-302
1172	PR00700B	Protein tyrosine phosphatase signature II	17.06 5.70e-21 313-333
1172	PR00700C	Protein tyrosine phosphatase signature III	13.89 9.05e-16 399-416
1172	PR00700D	Protein tyrosine phosphatase signature IV	12.83 7.84e-19 436-454
1172	PR00700E	Protein tyrosine phosphatase signature V	16.60 1.93e-10 467-482
1172	PR00700F	Protein tyrosine phosphatase signature VI	10.33 1.00e-11 483-493
1172	PR01371E	Salmonella/Yersinia modular tyrosine phosphatase signature V	11.71 2.11e-09 437-448
1187	IPB000008C	C2 domain	23.37 2.44e-09 61-100
1187	IPB002642A	Lysophospholipase catalytic domain	18.37 1.69e-13 355-380
1187	IPB002642B	Lysophospholipase catalytic domain	11.84 4.38e-15 383-407
1187	IPB002642E	Lysophospholipase catalytic domain	18.19 6.91e-10 509-534
1187	IPB002642G	Lysophospholipase catalytic domain	34.11 6.72e-10 678-726
1187	PR00360B	C2 domain signature II	11.64 8.67e-12 88-101
1188	PR00860B	Vertebrate metallothionein signature II	6.90 7.18e-09 52-65
1189	IPB000566B	Lipocalin and cytosolic fatty-acid binding protein	8.91 9.53e-09 122-132
1189	IPB002223	Pancreatic trypsin inhibitor (Kunitz) family	17.66 1.32e-23 285-319
1189	IPB002223	Pancreatic trypsin inhibitor (Kunitz) family	17.66 9.12e-15 247-281
1189	PR00179A	Lipocalin signature I	13.97 2.13e-10 38-50
1189	PR00179B	Lipocalin signature II	7.67 4.91e-12 122-134
1189	PR00179C	Lipocalin signature III	17.26 4.60e-09 150-165
1189	PR00759A	Basic protease (Kunitz-type) inhibitor family signature I	14.67 5.88e-09 228-242

Table 3B  
391

SEQ ID NO:	Database entry ID	Description	Results*
1189	PR00759B	Basic protease (Kunitz-type) inhibitor family signature II	12.35 3.81e-10 294-304
1189	PR00759B	Basic protease (Kunitz-type) inhibitor family signature II	12.35 8.71e-12 256-266
1189	PR00759C	Basic protease (Kunitz-type) inhibitor family signature III	12.43 6.00e-14 304-319
1189	PR01215A	Alpha-1-microglobulin signature I	14.08 3.25e-18 32-45
1189	PR01215B	Alpha-1-microglobulin signature II	14.28 7.00e-15 47-58
1189	PR01215C	Alpha-1-microglobulin signature III	8.22 3.63e-19 91-106
1189	PR01215D	Alpha-1-microglobulin signature IV	12.88 1.53e-27 113-132
1189	PR01215E	Alpha-1-microglobulin signature V	12.83 1.66e-24 145-166
1189	PR01215F	Alpha-1-microglobulin signature VI	10.31 9.59e-25 173-192
1189	PR01273A	Invertebrate colouration protein signature I	14.08 3.53e-09 33-49
1200	IPB001594	DHHC-type Zn-finger	23.81 7.47e-31 99-143
1201	IPB001909	KRAB box	17.37 9.61e-26 35-69
1202	IPB001909	KRAB box	17.37 9.61e-26 35-69
1203	IPB000961A	Protein kinase C-terminal domain	16.82 5.97e-09 74-108
1204	IPB000001D	Kringle	11.31 3.92e-15 81-97
1204	IPB000126A	Serine proteases, V8 family	11.75 9.51e-12 81-96
1204	IPB000177K	Apple domain	13.19 3.55e-11 83-115
1204	IPB001254A	Serine proteases, trypsin family	9.98 6.06e-18 81-97
1204	PR00722A	Chymotrypsin serine protease family (S1) signature I	12.06 3.37e-17 82-97
1208	PR00179B	Lipocalin signature II	7.67 7.30e-09 121-133
1208	PR01174F	Retinol binding protein signature VI	11.76 9.34e-09 120-136
1208	PR01254A	Prostaglandin D synthase signature I	12.32 9.79e-09 31-54
1208	PR01275B	Neutrophil gelatinase lipocalin signature II	9.02 3.16e-11 39-49
1209	IPB001007C	von Willebrand factor, type C repeat	20.89 5.97e-10 34-55
1212	PR00019A	Leucine-rich repeat signature I	11.72 2.50e-12 177-190
1212	PR00019A	Leucine-rich repeat signature I	11.72 7.82e-09 390-403
1212	PR00019B	Leucine-rich repeat signature II	11.42 1.64e-13 174-187
1212	PR00019B	Leucine-rich repeat signature II	11.42 9.18e-11 387-400
1213	IPB000834B	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	13.51 2.50e-17 103-117
1213	IPB000834C	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	17.20 2.80e-15 172-188
1213	IPB000834D	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	18.95 4.72e-12 199-225
1213	IPB000834E	Zinc carboxypeptidases, carboxypeptidase A metalloprotease (M14) family	9.80 2.15e-09 228-242
1213	PR00765B	Carboxypeptidase A metalloprotease (M14) family signature II	14.48 1.39e-15 99-113
1213	PR00765C	Carboxypeptidase A metalloprotease (M14) family signature III	10.88 1.82e-10 179-187
1213	PR00765D	Carboxypeptidase A metalloprotease (M14) family signature IV	14.06 9.45e-12 233-246
1214	PR01536C	Interleukin-1 receptor type I and type II family signature III	19.92 7.00e-09 67-90
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 1.30e-11 56-93
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.22e-15 44-81
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.29e-19 38-75
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.74e-09 20-57
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.82e-15 68-105
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 3.05e-10 92-129

Table 3B  
392

SEQ ID NO:	Database entry ID	Description	Results*
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 3.22e-11 26-63
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 3.34e-14 89-126
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 3.36e-12 32-69
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 3.57e-14 47-84
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 4.16e-15 71-108
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 4.17e-16 80-117
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 4.28e-15 41-78
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 4.62e-14 65-102
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 4.65e-15 50-87
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 5.56e-13 62-99
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 5.56e-14 83-120
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 5.91e-14 29-66
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 7.07e-10 95-132
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 7.46e-16 86-123
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 7.59e-16 77-114
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 8.29e-12 53-90
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 8.36e-14 59-96
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.00e-19 35-72
1221	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.42e-14 74-111
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.00e-20 72-125
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.22e-16 33-86
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.47e-18 51-104
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.51e-21 69-122
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.74e-20 36-89
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.94e-10 90-143
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.98e-09 15-68
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 2.21e-14 27-80
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 3.03e-17 78-131
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 3.22e-15 24-77
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 3.84e-13 84-137
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 3.95e-24 42-95
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 4.16e-18 30-83
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 4.32e-23 75-128
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 4.70e-20 60-113
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 5.07e-20 45-98
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.40e-09 3-56
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.61e-18 63-116
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.62e-22 57-110
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.63e-09 9-62
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.71e-09 12-65
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.81e-21 39-92
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.84e-19 66-119
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.95e-13 18-71
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 8.89e-10 93-146
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.07e-22 54-107
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.16e-12 81-134
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.41e-13 21-74
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.88e-20 48-101
1221	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.91e-12 87-140
1221	IPB001073A	Complement C1q protein	22.14 1.00e-11 40-74
1221	IPB001073A	Complement C1q protein	22.14 1.42e-09 55-89
1221	IPB001073A	Complement C1q protein	22.14 1.67e-11 79-113
1221	IPB001073A	Complement C1q protein	22.14 2.97e-09 43-77



Table 3B  
393

SEQ ID NO:	Database entry ID	Description	Results*
1221	IPB001073A	Complement C1q protein	22.14 3.39e-09 91-125
1221	IPB001073A	Complement C1q protein	22.14 3.75e-10 31-65
1221	IPB001073A	Complement C1q protein	22.14 4.09e-09 85-119
1221	IPB001073A	Complement C1q protein	22.14 4.45e-13 34-68
1221	IPB001073A	Complement C1q protein	22.14 4.60e-12 94-128
1221	IPB001073A	Complement C1q protein	22.14 5.50e-11 88-122
1221	IPB001073A	Complement C1q protein	22.14 5.67e-11 76-110
1221	IPB001073A	Complement C1q protein	22.14 7.10e-10 52-86
1221	IPB001073A	Complement C1q protein	22.14 7.33e-11 82-116
1221	IPB001073A	Complement C1q protein	22.14 7.51e-13 37-71
1221	IPB001073A	Complement C1q protein	22.14 7.83e-11 97-131
1221	IPB001073A	Complement C1q protein	22.14 8.17e-10 58-92
1221	IPB001073A	Complement C1q protein	22.14 8.59e-09 49-83
1221	IPB001073A	Complement C1q protein	22.14 9.81e-13 100-134
1221	IPB001073B	Complement C1q protein	20.88 7.75e-28 150-184
1221	IPB001073C	Complement C1q protein	13.07 8.56e-16 214-233
1221	IPB001073D	Complement C1q protein	7.60 7.68e-09 251-260
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 1.09e-13 82-134
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 1.39e-09 91-143
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 1.42e-16 64-116
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.41e-17 31-83
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.47e-16 76-128
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.73e-17 67-119
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.89e-20 49-101
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.97e-09 10-62
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.03e-16 79-131
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.79e-20 46-98
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.80e-17 25-77
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.80e-17 28-80
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.82e-21 40-92
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.96e-12 13-65
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 5.85e-09 88-140
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.56e-18 34-86
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.77e-14 22-74
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 7.39e-23 37-89
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 7.55e-14 16-68

Table 3B  
394

SEQ ID NO:	Database entry ID	Description	Results*
		4 procollagen	
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 7.99e-20 70-122
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 8.07e-22 43-95
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 8.30e-23 61-113
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 8.58e-20 58-110
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 9.09e-15 19-71
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 9.18e-19 55-107
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 9.37e-12 85-137
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 9.56e-18 73-125
1221	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 9.59e-24 52-104
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 1.00e-11 104-124
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 1.49e-09 47-67
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 1.85e-09 107-127
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 2.82e-09 41-61
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 2.82e-09 86-106
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 3.31e-09 71-91
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 3.55e-09 56-76
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 3.68e-09 50-70
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 3.92e-09 68-88
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 5.14e-09 89-109
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 5.24e-10 101-121
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 6.01e-10 38-58
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 6.72e-09 59-79
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 7.81e-09 35-55
1221	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 9.15e-09 44-64
1221	PR00007A	Complement C1Q domain signature I	20.64 7.47e-17 143-169
1221	PR00007B	Complement C1Q domain signature II	15.63 5.26e-14 170-189
1221	PR00007C	Complement C1Q domain signature III	16.13 2.07e-14 214-235
1221	PR00007D	Complement C1Q domain signature IV	9.66 3.91e-10 249-259
1223	IPB000053H	Thymidine/pyrimidine-nucleoside	12.28 9.76e-09 28-37

Table 3B  
395

SEQ ID NO:	Database entry ID	Description	Results*
		phosphorylase	
1223	IPB000095E	PAK-box /P21-Rho-binding	17.62 4.19e-11 275-320
1223	IPB000494C	Epidermal growth-factor receptor (EGFR), L domain	24.40 3.60e-31 261-307
1223	IPB000494D	Epidermal growth-factor receptor (EGFR), L domain	19.44 7.50e-29 322-371
1223	IPB000494E	Epidermal growth-factor receptor (EGFR), L domain	21.65 2.46e-24 379-414
1223	IPB000861F	PKN/rhopilin/rhotekin rho-binding repeat	16.50 3.15e-11 268-322
1223	IPB000959B	POLO box duplicated region	15.68 1.99e-10 251-291
1223	IPB000961C	Protein kinase C-terminal domain	15.48 6.29e-13 274-308
1223	IPB001245A	Tyrosine kinase catalytic domain	22.45 7.19e-26 267-307
1223	IPB001245B	Tyrosine kinase catalytic domain	21.68 5.24e-23 325-363
1223	IPB001359H	Synapsin	22.58 3.51e-10 803-853
1223	IPB001359H	Synapsin	22.58 6.85e-09 777-827
1223	IPB001772C	Kinase associated domain 1	20.66 7.93e-13 262-292
1223	IPB003306E	WIF domain	25.51 5.35e-12 379-424
1223	IPB003527C	MAP kinase	14.70 2.48e-14 259-307
1223	PR00109A	Tyrosine kinase catalytic domain signature I	12.56 5.14e-10 237-250
1223	PR00109B	Tyrosine kinase catalytic domain signature II	11.07 2.61e-18 274-292
1223	PR00109C	Tyrosine kinase catalytic domain signature III	11.86 1.87e-10 324-334
1223	PR00109D	Tyrosine kinase catalytic domain signature IV	17.61 7.12e-21 343-365
1223	PR00109E	Tyrosine kinase catalytic domain signature V	12.99 5.09e-20 388-410
1223	PR01217B	Proline rich extensin signature II	4.82 6.09e-09 927-943
1224	PR00453A	Von Willebrand factor type A domain signature I	11.78 3.15e-09 50-67
1225	IPB000734	Lipase	10.25 8.13e-09 164-178
1227	IPB001304A	C-type lectin domain	17.98 3.48e-10 41-65
1227	IPB001304B	C-type lectin domain	10.18 5.95e-09 99-111
1227	PR01504F	Pancreatitis-associated protein signature VI	12.72 6.45e-09 153-167
1237	IPB002181A	Fibrinogen beta and gamma chains C-terminal globular domain	18.44 2.32e-10 244-260
1237	IPB002181B	Fibrinogen beta and gamma chains C-terminal globular domain	20.16 9.57e-10 271-307
1237	IPB002181C	Fibrinogen beta and gamma chains C-terminal globular domain	15.87 1.00e-08 294-306
1237	IPB002181D	Fibrinogen beta and gamma chains C-terminal globular domain	29.18 5.14e-19 312-352
1257	IPB000483	Leucine rich repeat C-terminal domain	11.18 1.58e-09 423-437
1257	PR00019B	Leucine-rich repeat signature II	11.42 8.33e-10 124-137
1257	PR00019B	Leucine-rich repeat signature II	11.42 8.33e-10 316-329
1263	IPB001762A	Disintegrin	23.93 6.45e-17 434-474
1263	IPB001762A	Disintegrin	23.93 6.47e-09 416-456
1263	IPB002870A	Reprolysin family propeptide	12.22 3.31e-10 63-79
1263	IPB002870B	Reprolysin family propeptide	24.73 6.56e-14 109-147
1263	IPB002870E	Reprolysin family propeptide	11.90 2.50e-11 341-353
1263	PR00289A	Disintegrin signature I	14.29 2.65e-13 450-469
1263	PR00873D	Echinoidea (sea urchin) metallothionein signature IV	8.25 7.35e-09 608-626
1264	IPB002168A	Lipolytic enzymes	11.40 5.07e-11 139-153

Table 3B  
396

SEQ ID NO:	Database entry ID	Description	Results*
1264	IPB002168B	Lipolytic enzymes	30.88 5.14e-19 157-204
1264	IPB002168C	Lipolytic enzymes	9.77 4.00e-16 216-229
1264	IPB002168D	Lipolytic enzymes	19.85 6.70e-18 369-393
1264	IPB002469I	Dipeptidyl peptidase IV, N-terminus	10.99 6.29e-09 211-229
1264	PR01040B	Tyrosyl-tRNA synthetase signature II	14.69 7.00e-09 245-260
1266	IPB000753A	Clusterin	12.20 6.45e-18 2-16
1266	IPB000753B	Clusterin	11.42 1.00e-40 39-83
1266	IPB000753D	Clusterin	15.34 1.00e-40 163-204
1266	IPB000753E	Clusterin	13.24 1.00e-40 230-279
1266	IPB000753F	Clusterin	17.89 1.00e-40 310-361
1266	IPB000753G	Clusterin	17.73 9.55e-24 372-390
1267	IPB000215A	Serpins	13.01 9.14e-18 107-130
1267	IPB000215C	Serpins	13.90 7.88e-15 211-225
1267	IPB000215D	Serpins	15.35 3.74e-17 314-340
1267	IPB000215E	Serpins	15.36 6.68e-16 401-425
1268	IPB000215A	Serpins	13.01 9.14e-18 125-148
1268	IPB000215B	Serpins	9.87 1.82e-11 218-230
1268	IPB000215C	Serpins	13.90 7.88e-15 247-261
1268	IPB000215D	Serpins	15.35 3.74e-17 350-376
1268	IPB000215E	Serpins	15.36 7.19e-18 437-461
1273	IPB001307A	Rhodanese signatures	11.98 8.64e-10 10-20
1273	IPB001307B	Rhodanese signatures	16.09 3.89e-14 48-61
1273	IPB001307C	Rhodanese signatures	15.11 1.28e-15 82-101
1273	IPB001307D	Rhodanese signatures	15.87 4.00e-26 110-135
1273	IPB001307E	Rhodanese signatures	13.19 5.36e-10 180-191
1273	IPB001307G	Rhodanese signatures	11.73 9.31e-16 251-263
1274	IPB000001D	Kringle	11.31 4.95e-14 147-163
1274	IPB000001G	Kringle	29.29 4.46e-11 259-287
1274	IPB000001H	Kringle	12.24 2.80e-10 299-309
1274	IPB000126A	Serine proteases, V8 family	11.75 9.25e-10 147-162
1274	IPB000177K	Apple domain	13.19 2.18e-11 149-181
1274	IPB000177L	Apple domain	11.33 9.92e-10 178-216
1274	IPB000177N	Apple domain	10.17 5.19e-24 289-323
1274	IPB000177O	Apple domain	14.39 9.00e-16 324-352
1274	IPB001254A	Serine proteases, trypsin family	9.98 3.37e-16 147-163
1274	IPB001254B	Serine proteases, trypsin family	15.01 5.50e-16 297-320
1274	IPB001254C	Serine proteases, trypsin family	16.54 4.86e-20 327-346
1274	IPB002093D	BRCA2 repeat	8.99 9.77e-09 8-46
1274	PR00722A	Chymotrypsin serine protease family (S1) signature I	12.06 3.88e-15 148-163
1274	PR00722B	Chymotrypsin serine protease family (S1) signature II	12.69 3.74e-09 203-217
1274	PR00722C	Chymotrypsin serine protease family (S1) signature III	10.74 1.69e-14 296-308
1279	PR00704I	Calpain cysteine protease (C2) family signature IX	9.42 4.14e-14 145-173
1284	IPB001706	Ribosomal protein L35	21.01 1.00e-18 119-157
1285	IPB001706	Ribosomal protein L35	21.01 1.00e-18 119-157
1287	IPB000822	Zinc finger, C2H2 type	14.67 1.43e-11 399-424
1287	IPB000822	Zinc finger, C2H2 type	14.67 2.50e-20 259-284
1287	IPB000822	Zinc finger, C2H2 type	14.67 3.25e-19 231-256
1287	IPB000822	Zinc finger, C2H2 type	14.67 5.50e-19 343-368
1287	IPB000822	Zinc finger, C2H2 type	14.67 6.54e-18 203-228

Table 3B  
397

SEQ ID NO:	Database entry ID	Description	Results*
1287	IPB000822	Zinc finger, C2H2 type	14.67 8.50e-13 315-340
1287	IPB000822	Zinc finger, C2H2 type	14.67 8.62e-18 287-312
1287	IPB000822	Zinc finger, C2H2 type	14.67 9.10e-22 175-200
1287	IPB000822	Zinc finger, C2H2 type	14.67 9.10e-22 371-396
1287	IPB001275	DM DNA binding domain	19.17 2.75e-12 359-398
1287	IPB001275	DM DNA binding domain	19.17 3.57e-11 275-314
1287	IPB001275	DM DNA binding domain	19.17 4.34e-11 191-230
1287	IPB001275	DM DNA binding domain	19.17 9.60e-12 247-286
1287	IPB001275	DM DNA binding domain	19.17 9.87e-12 331-370
1287	IPB001909	KRAB box	17.37 3.33e-22 16-50
1287	PR00048A	C2H2-type zinc finger signature I	9.94 2.59e-12 396-409
1287	PR00048A	C2H2-type zinc finger signature I	9.94 2.88e-09 340-353
1287	PR00048A	C2H2-type zinc finger signature I	9.94 3.77e-14 368-381
1287	PR00048A	C2H2-type zinc finger signature I	9.94 5.15e-14 284-297
1287	PR00048A	C2H2-type zinc finger signature I	9.94 6.63e-09 312-325
1287	PR00048A	C2H2-type zinc finger signature I	9.94 6.73e-10 256-269
1287	PR00048A	C2H2-type zinc finger signature I	9.94 9.36e-13 200-213
1287	PR00048B	C2H2-type zinc finger signature II	5.52 1.00e-12 188-197
1287	PR00048B	C2H2-type zinc finger signature II	5.52 3.25e-11 384-393
1287	PR00048B	C2H2-type zinc finger signature II	5.52 4.00e-09 356-365
1287	PR00048B	C2H2-type zinc finger signature II	5.52 4.86e-10 244-253
1287	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-09 300-309
1287	PR00048B	C2H2-type zinc finger signature II	5.52 7.00e-12 272-281
1292	IPB000034C	Laminin B	12.97 7.31e-09 54-72
1292	IPB000561	EGF-like domain	4.89 8.07e-09 54-62
1292	IPB001774D	Delta serrate ligand	19.23 5.91e-09 20-66
1294	IPB001007B	von Willebrand factor, type C repeat	10.03 2.89e-09 92-101
1295	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-18 433-458
1295	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-22 405-430
1295	IPB000822	Zinc finger, C2H2 type	14.67 1.90e-22 321-346
1295	IPB000822	Zinc finger, C2H2 type	14.67 4.00e-10 265-290
1295	IPB000822	Zinc finger, C2H2 type	14.67 4.86e-10 489-514
1295	IPB000822	Zinc finger, C2H2 type	14.67 6.25e-19 377-402
1295	IPB000822	Zinc finger, C2H2 type	14.67 7.75e-20 461-486
1295	IPB000822	Zinc finger, C2H2 type	14.67 8.50e-19 349-374
1295	IPB001222	TFIIS zinc ribbon domain	24.63 3.66e-09 433-469
1295	IPB001275	DM DNA binding domain	19.17 3.85e-09 309-348
1295	IPB001275	DM DNA binding domain	19.17 4.77e-10 337-376
1295	IPB001275	DM DNA binding domain	19.17 5.22e-09 421-460
1295	IPB001275	DM DNA binding domain	19.17 8.42e-10 393-432
1295	IPB001275	DM DNA binding domain	19.17 9.27e-10 449-488
1295	IPB001909	KRAB box	17.37 9.50e-32 10-44
1295	PR00048A	C2H2-type zinc finger signature I	9.94 1.00e-10 318-331
1295	PR00048A	C2H2-type zinc finger signature I	9.94 1.00e-13 402-415
1295	PR00048A	C2H2-type zinc finger signature I	9.94 1.64e-13 346-359
1295	PR00048A	C2H2-type zinc finger signature I	9.94 5.76e-12 374-387
1295	PR00048A	C2H2-type zinc finger signature I	9.94 5.85e-14 458-471
1295	PR00048A	C2H2-type zinc finger signature I	9.94 7.92e-14 486-499
1295	PR00048A	C2H2-type zinc finger signature I	9.94 8.62e-14 430-443
1295	PR00048B	C2H2-type zinc finger signature II	5.52 1.50e-09 418-427
1295	PR00048B	C2H2-type zinc finger signature II	5.52 3.50e-09 474-483
1295	PR00048B	C2H2-type zinc finger signature II	5.52 6.79e-10 334-343
1295	PR00048B	C2H2-type zinc finger signature II	5.52 8.50e-09 390-399

Table 3B  
398

SEQ ID NO:	Database entry ID	Description	Results*
1296	IPB000822	Zinc finger, C2H2 type	14.67 1.69e-18 304-329
1296	IPB000822	Zinc finger, C2H2 type	14.67 3.25e-14 444-469
1296	IPB000822	Zinc finger, C2H2 type	14.67 3.77e-18 220-245
1296	IPB000822	Zinc finger, C2H2 type	14.67 6.25e-19 248-273
1296	IPB000822	Zinc finger, C2H2 type	14.67 6.73e-21 416-441
1296	IPB000822	Zinc finger, C2H2 type	14.67 7.00e-20 276-301
1296	IPB000822	Zinc finger, C2H2 type	14.67 8.20e-16 388-413
1296	IPB000822	Zinc finger, C2H2 type	14.67 8.20e-23 192-217
1296	IPB000822	Zinc finger, C2H2 type	14.67 8.36e-21 332-357
1296	IPB000822	Zinc finger, C2H2 type	14.67 8.50e-19 472-497
1296	IPB000822	Zinc finger, C2H2 type	14.67 8.62e-18 500-525
1296	IPB000822	Zinc finger, C2H2 type	14.67 9.18e-21 360-385
1296	IPB001275	DM DNA binding domain	19.17 1.97e-10 488-527
1296	IPB001275	DM DNA binding domain	19.17 2.37e-09 376-415
1296	IPB001275	DM DNA binding domain	19.17 2.95e-10 264-303
1296	IPB001275	DM DNA binding domain	19.17 3.96e-09 236-275
1296	IPB001275	DM DNA binding domain	19.17 5.10e-09 432-471
1296	IPB001275	DM DNA binding domain	19.17 5.38e-10 208-247
1296	IPB001275	DM DNA binding domain	19.17 5.50e-10 404-443
1296	IPB001275	DM DNA binding domain	19.17 5.76e-11 460-499
1296	IPB001275	DM DNA binding domain	19.17 5.90e-09 180-219
1296	IPB001275	DM DNA binding domain	19.17 6.79e-11 348-387
1296	IPB001275	DM DNA binding domain	19.17 8.30e-10 320-359
1296	IPB001909	KRAB box	17.37 5.50e-37 16-50
1296	PR00048A	C2H2-type zinc finger signature I	9.94 1.38e-09 217-230
1296	PR00048A	C2H2-type zinc finger signature I	9.94 2.06e-12 245-258
1296	PR00048A	C2H2-type zinc finger signature I	9.94 2.06e-12 301-314
1296	PR00048A	C2H2-type zinc finger signature I	9.94 3.86e-10 273-286
1296	PR00048A	C2H2-type zinc finger signature I	9.94 4.71e-12 413-426
1296	PR00048A	C2H2-type zinc finger signature I	9.94 4.86e-13 329-342
1296	PR00048A	C2H2-type zinc finger signature I	9.94 5.50e-15 385-398
1296	PR00048A	C2H2-type zinc finger signature I	9.94 5.50e-15 469-482
1296	PR00048A	C2H2-type zinc finger signature I	9.94 5.91e-10 497-510
1296	PR00048A	C2H2-type zinc finger signature I	9.94 6.14e-13 189-202
1296	PR00048A	C2H2-type zinc finger signature I	9.94 6.14e-13 525-538
1296	PR00048A	C2H2-type zinc finger signature I	9.94 6.82e-12 357-370
1296	PR00048B	C2H2-type zinc finger signature II	5.52 1.00e-10 261-270
1296	PR00048B	C2H2-type zinc finger signature II	5.52 1.00e-10 289-298
1296	PR00048B	C2H2-type zinc finger signature II	5.52 2.29e-10 205-214
1296	PR00048B	C2H2-type zinc finger signature II	5.52 4.86e-10 513-522
1296	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-10 317-326
1296	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-10 485-494
1296	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-11 345-354
1296	PR00048B	C2H2-type zinc finger signature II	5.52 6.79e-10 233-242
1296	PR00048B	C2H2-type zinc finger signature II	5.52 7.75e-11 429-438
1299	IPB000130	Neutral zinc metalloproteases, zinc-binding region	5.86 1.00e-11 188-198
1299	IPB001818B	Matrixin	26.48 4.63e-13 87-128
1299	IPB001818C	Matrixin	24.38 4.54e-09 134-179
1299	IPB001818D	Matrixin	14.91 3.36e-25 182-213
1299	IPB001818E	Matrixin	8.86 4.26e-10 222-235
1299	IPB002870E	Reprolysins family propeptide	11.90 2.00e-09 190-202
1299	PR00138A	Matrixin signature I	12.54 9.25e-10 37-50

Table 3B  
399

SEQ ID NO:	Database entry ID	Description	Results*
1299	PR00138B	Matrixin signature II	14.84 4.51e-11 106-121
1299	PR00138C	Matrixin signature III	20.07 8.86e-11 132-160
1299	PR00138D	Matrixin signature IV	14.57 1.86e-21 188-213
1299	PR00138E	Matrixin signature V	7.10 9.77e-11 222-235
1299	PR00480B	Astacin family signature II	14.35 1.14e-09 183-201
1301	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 6.09e-11 150-187
1305	IPB002579C	Domain of unknown function DUF25	13.50 3.89e-28 75-102
1305	IPB002579D	Domain of unknown function DUF25	11.55 7.92e-12 128-139
1305	IPB002579E	Domain of unknown function DUF25	12.14 9.31e-13 149-159
1307	IPB002350	Kazal-type serine protease inhibitor family	31.78 4.09e-16 28-68
1307	PR00290B	Kazal-type serine protease inhibitor signature II	16.63 1.26e-10 39-50
1308	IPB001462G	Hepadnaviral P protein C-terminal domain	17.22 3.38e-09 52-72
1309	IPB001271	Mammalian defensin	19.97 3.42e-21 139-167
1312	IPB002041A	GTP-binding nuclear protein Ran family	11.12 9.36e-12 63-107
1312	PR00449A	Transforming protein P21 ras signature I	12.48 2.57e-16 63-84
1312	PR00449B	Transforming protein P21 ras signature II	14.39 4.81e-09 86-102
1312	PR00449C	Transforming protein P21 ras signature III	15.70 8.26e-12 103-125
1312	PR00449D	Transforming protein P21 ras signature IV	10.27 6.50e-13 167-180
1312	PR00449E	Transforming protein P21 ras signature V	13.39 2.04e-13 201-223
1314	IPB000822	Zinc finger, C2H2 type	14.67 1.69e-18 165-190
1314	IPB000822	Zinc finger, C2H2 type	14.67 1.75e-19 305-330
1314	IPB000822	Zinc finger, C2H2 type	14.67 2.93e-17 277-302
1314	IPB000822	Zinc finger, C2H2 type	14.67 3.08e-18 193-218
1314	IPB000822	Zinc finger, C2H2 type	14.67 6.40e-16 221-246
1314	IPB000822	Zinc finger, C2H2 type	14.67 6.63e-14 333-358
1314	IPB000822	Zinc finger, C2H2 type	14.67 9.31e-18 249-274
1314	IPB001275	DM DNA binding domain	19.17 2.88e-12 321-360
1314	IPB001275	DM DNA binding domain	19.17 3.43e-10 209-248
1314	IPB001275	DM DNA binding domain	19.17 4.16e-10 265-304
1314	IPB001275	DM DNA binding domain	19.17 4.86e-11 237-276
1314	IPB001275	DM DNA binding domain	19.17 5.26e-10 293-332
1314	IPB001275	DM DNA binding domain	19.17 6.66e-11 181-220
1314	PR00048A	C2H2-type zinc finger signature I	9.94 2.23e-10 302-315
1314	PR00048A	C2H2-type zinc finger signature I	9.94 2.29e-13 330-343
1314	PR00048A	C2H2-type zinc finger signature I	9.94 2.93e-13 218-231
1314	PR00048A	C2H2-type zinc finger signature I	9.94 4.86e-13 190-203
1314	PR00048A	C2H2-type zinc finger signature I	9.94 5.26e-11 274-287
1314	PR00048A	C2H2-type zinc finger signature I	9.94 6.14e-13 162-175
1314	PR00048A	C2H2-type zinc finger signature I	9.94 8.07e-13 358-371
1314	PR00048A	C2H2-type zinc finger signature I	9.94 8.88e-09 246-259
1314	PR00048B	C2H2-type zinc finger signature II	5.52 2.29e-10 178-187
1314	PR00048B	C2H2-type zinc finger signature II	5.52 3.57e-10 318-327
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 1.58e-10 78-123
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 1.92e-09 88-133
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 2.06e-10 63-108
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 2.65e-09 28-73
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 2.93e-09 5-50
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 3.07e-12 72-117
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 3.57e-09 82-127
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 4.86e-09 83-128
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 5.06e-10 56-101

Table 3B  
400

SEQ ID NO:	Database entry ID	Description	Results*
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 5.06e-10 77-122
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 5.10e-11 46-91
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 5.68e-09 36-81
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 5.96e-09 48-93
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 6.30e-11 87-132
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 6.33e-09 25-70
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 6.97e-09 93-138
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 7.48e-10 15-60
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 8.20e-11 101-146
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 8.44e-09 73-118
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 8.81e-09 61-106
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 8.84e-10 33-78
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 9.38e-12 34-79
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 9.82e-09 91-136
1315	IPB000006	Vertebrate metallothionein, family 1	13.41 9.90e-10 102-147
1315	IPB000359B	Cystine-knot domain	19.26 3.48e-10 129-147
1315	IPB000359B	Cystine-knot domain	19.26 6.34e-09 90-108
1315	IPB000437	Prokaryotic membrane lipoprotein lipid attachment site	13.83 7.23e-10 10-21
1315	IPB000437	Prokaryotic membrane lipoprotein lipid attachment site	13.83 9.31e-10 20-31
1315	IPB000726B	Glycoside hydrolase family 19	5.39 5.68e-11 8-18
1315	IPB000726B	Glycoside hydrolase family 19	5.39 7.84e-11 18-28
1315	IPB001007C	von Willebrand factor, type C repeat	20.89 4.27e-09 74-95
1315	IPB001169K	Integrin beta, C-terminus	27.45 1.00e-10 10-52
1315	IPB001169K	Integrin beta, C-terminus	27.45 2.21e-11 81-123
1315	IPB001169K	Integrin beta, C-terminus	27.45 4.04e-09 18-60
1315	IPB001169K	Integrin beta, C-terminus	27.45 4.77e-13 3-45
1315	IPB001169K	Integrin beta, C-terminus	27.45 7.31e-11 13-55
1315	IPB001271	Mammalian defensin	19.97 1.00e-10 12-40
1315	IPB001762A	Disintegrin	23.93 2.24e-09 75-115
1315	IPB001762A	Disintegrin	23.93 2.41e-09 95-135
1315	IPB001762A	Disintegrin	23.93 5.41e-09 114-154
1315	IPB002174A	Furin-like cysteine rich region	30.51 1.79e-11 5-36
1315	IPB002174A	Furin-like cysteine rich region	30.51 1.90e-09 56-87
1315	IPB002174A	Furin-like cysteine rich region	30.51 3.19e-10 12-43
1315	IPB002174A	Furin-like cysteine rich region	30.51 9.16e-12 2-33
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.00e-08 15-48
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.11e-11 122-155
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.67e-12 68-101
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.72e-14 74-107
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.75e-11 49-82
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.00e-12 11-44
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.07e-11 64-97
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.33e-12 54-87
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.85e-17 78-111
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.42e-09 98-131
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.52e-09 79-112
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.00e-15 117-150
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.60e-10 21-54
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.65e-09 102-135
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.93e-11 5-38
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.00e-15 83-116



Table 3B  
401

SEQ ID NO:	Database entry ID	Description	Results*
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.01e-10 63-96
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.13e-09 6-39
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.56e-17 73-106
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.77e-13 16-49
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.24e-10 118-151
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.67e-12 113-146
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.00e-19 112-145
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.16e-10 59-92
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.56e-14 93-126
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.57e-19 88-121
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.65e-09 69-102
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.67e-10 84-117
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.03e-09 89-122
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.71e-09 10-43
1315	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.80e-10 103-136
1315	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.27e-10 141-155
1315	IPB002494B	Keratin, high sulfur B2 protein	10.58 3.92e-09 107-121
1315	IPB002494B	Keratin, high sulfur B2 protein	10.58 4.49e-10 131-145
1315	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.10e-12 102-116
1315	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.10e-12 92-106
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.00e-08 19-62
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.52e-15 86-129
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.59e-09 14-57
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.79e-10 8-51
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.55e-11 62-105
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.66e-13 61-104
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.39e-17 101-144
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.46e-16 110-153
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.80e-10 80-123
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.71e-25 81-124
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.76e-10 67-110
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.11e-10 111-154
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.11e-10 52-95
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.84e-18 76-119
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.42e-10 51-94
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.00e-15 105-148
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.47e-10 85-128
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.62e-15 96-139
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.64e-16 66-109
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.98e-09 95-138
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.07e-24 91-134
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.13e-12 106-149
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.44e-21 71-114
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.45e-11 72-115
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.70e-14 90-133
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.00e-11 46-89
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.30e-14 100-143
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.65e-10 75-118
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.82e-11 77-120
1315	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.83e-09 39-82
1315	IPB003571B	Snake toxin	18.08 1.20e-09 22-45
1315	IPB003854B	Gibberellin regulated protein	19.26 9.51e-10 45-83
1315	PR00858B	Crustacean metallothionein signature II	5.93 3.39e-09 49-67

Table 3B  
402

SEQ ID NO:	Database entry ID	Description	Results*
1315	PR00858B	Crustacean metallothionein signature II	5.93 7.22e-09 15-33
1315	PR00874C	Fungi-IV metallothionein signature III	4.37 3.75e-09 48-62
1315	PR00876B	Nematode metallothionein signature II	7.66 5.01e-09 89-102
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 1.29e-10 57-102
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 4.12e-09 37-82
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 5.41e-09 52-97
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 5.59e-09 38-83
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 5.68e-09 68-113
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 5.70e-11 47-92
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 5.74e-10 87-132
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 5.87e-09 18-63
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 6.20e-11 88-133
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 6.60e-09 17-62
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 8.35e-09 23-68
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 8.35e-09 78-123
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 8.45e-10 92-137
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 8.70e-11 22-67
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 8.99e-09 33-78
1316	IPB000006	Vertebrate metallothionein, family 1	13.41 9.45e-09 43-88
1316	IPB000359B	Cystine-knot domain	19.26 1.00e-14 90-108
1316	IPB000359B	Cystine-knot domain	19.26 1.56e-09 45-63
1316	IPB000359B	Cystine-knot domain	19.26 2.97e-09 80-98
1316	IPB000359B	Cystine-knot domain	19.26 5.78e-09 120-138
1316	IPB000359B	Cystine-knot domain	19.26 6.25e-12 20-38
1316	IPB000359B	Cystine-knot domain	19.26 8.31e-09 55-73
1316	IPB000359B	Cystine-knot domain	19.26 9.16e-09 25-43
1316	IPB000359B	Cystine-knot domain	19.26 9.16e-09 95-113
1316	IPB000726B	Glycoside hydrolase family 19	5.39 8.46e-09 21-31
1316	IPB000970F	Developmental signaling protein, Wnt-1 family	23.43 5.94e-09 59-107
1316	IPB001271	Mammalian defensin	19.97 1.35e-10 25-53
1316	IPB001271	Mammalian defensin	19.97 1.88e-10 94-122
1316	IPB001271	Mammalian defensin	19.97 2.94e-10 75-103
1316	IPB001271	Mammalian defensin	19.97 4.93e-09 50-78
1316	IPB001271	Mammalian defensin	19.97 4.93e-09 95-123
1316	IPB001271	Mammalian defensin	19.97 5.24e-10 24-52
1316	IPB001271	Mammalian defensin	19.97 5.25e-09 30-58
1316	IPB001271	Mammalian defensin	19.97 5.58e-09 85-113
1316	IPB001271	Mammalian defensin	19.97 7.18e-10 80-108
1316	IPB001271	Mammalian defensin	19.97 7.22e-09 65-93
1316	IPB001271	Mammalian defensin	19.97 7.35e-10 40-68
1316	IPB001271	Mammalian defensin	19.97 8.06e-10 89-117
1316	IPB001271	Mammalian defensin	19.97 8.20e-09 35-63
1316	IPB001271	Mammalian defensin	19.97 8.31e-11 15-43
1316	IPB001271	Mammalian defensin	19.97 8.36e-09 84-112
1316	IPB001271	Mammalian defensin	19.97 8.59e-10 39-67
1316	IPB001271	Mammalian defensin	19.97 8.85e-09 64-92
1316	IPB001271	Mammalian defensin	19.97 8.94e-10 49-77
1316	IPB001271	Mammalian defensin	19.97 9.35e-09 74-102
1316	IPB001271	Mammalian defensin	19.97 9.67e-09 34-62
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.97e-09 94-127
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.23e-18 53-86
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.85e-13 108-141

Table 3B  
403

SEQ ID NO:	Database entry ID	Description	Results*
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.86e-19 48-81
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.94e-09 49-82
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.33e-12 3-36
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.38e-15 13-46
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.44e-12 18-51
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.67e-12 8-41
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.71e-19 93-126
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.48e-09 54-87
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.68e-10 118-151
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.71e-17 28-61
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.89e-10 24-57
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.97e-09 64-97
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.99e-10 29-62
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.04e-13 43-76
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.10e-17 38-71
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.84e-09 59-92
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.92e-14 88-121
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.63e-15 83-116
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.77e-13 68-101
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.22e-17 73-106
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.58e-09 19-52
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.63e-15 103-136
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.82e-18 23-56
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.07e-16 58-91
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.33e-16 63-96
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.94e-17 98-131
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.03e-09 39-72
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.16e-14 78-111
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.69e-10 89-122
1316	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.76e-14 33-66
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.28e-11 22-36
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 2.27e-09 87-101
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 3.69e-10 92-106
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.22e-11 77-91
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.78e-11 72-86
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.07e-09 52-66
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.20e-11 37-51
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.64e-10 27-41
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.64e-10 97-111
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.91e-10 107-121
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.91e-10 62-76
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 7.08e-09 112-126
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 7.08e-09 67-81
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.25e-12 47-61
1316	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.44e-11 117-131
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.00e-08 15-58
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.18e-11 55-98
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.22e-17 101-144
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.36e-11 95-138
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.54e-16 1-44
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.60e-31 41-84
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.94e-26 26-69
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.76e-13 25-68

Table 3B  
404

SEQ ID NO:	Database entry ID	Description	Results*
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.45e-10 20-63
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.47e-22 56-99
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.09e-12 45-88
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.19e-12 116-159
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.26e-18 21-64
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.47e-24 81-124
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.69e-17 6-49
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.80e-28 76-119
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.84e-10 65-108
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.93e-10 60-103
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.95e-09 100-143
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.01e-13 90-133
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.50e-25 31-74
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.54e-10 35-78
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.54e-10 80-123
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.42e-10 121-164
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.50e-23 46-89
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.52e-21 91-134
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.64e-26 96-139
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.91e-11 85-128
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.04e-22 111-154
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.12e-10 70-113
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.67e-33 16-59
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.75e-28 86-129
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.34e-10 75-118
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.53e-13 50-93
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.70e-20 11-54
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.74e-09 10-53
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.88e-23 36-79
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.88e-28 66-109
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.30e-10 30-73
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.32e-27 51-94
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.38e-23 71-114
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.57e-17 106-149
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.58e-09 105-148
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.60e-25 61-104
1316	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.91e-10 40-83
1316	IPB003888D	FY-rich domain N-terminus	24.70 6.16e-09 100-142
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 1.00e-08 38-83
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 1.18e-09 165-210
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 1.28e-09 85-130
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 1.39e-10 130-175
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 1.46e-09 60-105
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 1.55e-09 70-115
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 2.29e-09 176-221
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 2.35e-10 64-109
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 3.11e-09 69-114
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 3.32e-10 104-149
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 3.94e-09 239-284
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 4.12e-09 125-170
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.32e-09 201-246
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.59e-09 190-235
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.68e-09 175-220

Table 3B  
405

SEQ ID NO:	Database entry ID	Description	Results*
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.68e-09 19-64
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.84e-10 210-255
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.87e-09 80-125
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.94e-10 170-215
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 5.94e-10 200-245
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 6.23e-09 186-231
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 6.88e-09 160-205
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 7.39e-10 225-270
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 7.43e-09 115-160
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 7.43e-09 55-100
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 7.68e-10 24-69
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 7.97e-10 215-260
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 8.35e-09 15-60
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 8.35e-09 45-90
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 8.62e-09 95-140
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 9.13e-10 120-165
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 9.27e-09 205-250
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 9.27e-09 206-251
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 9.45e-09 126-171
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 9.61e-10 34-79
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 9.91e-09 151-196
1317	IPB000006	Vertebrate metallothionein, family 1	13.41 9.91e-09 29-74
1317	IPB000359B	Cystine-knot domain	19.26 1.28e-09 128-146
1317	IPB000359B	Cystine-knot domain	19.26 2.80e-11 163-181
1317	IPB000359B	Cystine-knot domain	19.26 3.53e-09 243-261
1317	IPB000359B	Cystine-knot domain	19.26 4.38e-09 22-40
1317	IPB000359B	Cystine-knot domain	19.26 6.06e-09 248-266
1317	IPB000359B	Cystine-knot domain	19.26 6.34e-09 203-221
1317	IPB000359B	Cystine-knot domain	19.26 6.63e-12 198-216
1317	IPB000359B	Cystine-knot domain	19.26 7.75e-12 62-80
1317	IPB000359B	Cystine-knot domain	19.26 8.88e-09 72-90
1317	IPB000967E	Zinc finger NF-X1 type	21.88 3.73e-09 111-151
1317	IPB000967E	Zinc finger NF-X1 type	21.88 6.46e-09 89-129
1317	IPB001169K	Integrin beta, C-terminus	27.45 5.38e-09 216-258
1317	IPB001271	Mammalian defensin	19.97 1.00e-08 61-89
1317	IPB001271	Mammalian defensin	19.97 1.49e-09 132-160
1317	IPB001271	Mammalian defensin	19.97 1.88e-10 42-70
1317	IPB001271	Mammalian defensin	19.97 2.15e-09 193-221
1317	IPB001271	Mammalian defensin	19.97 2.31e-09 163-191
1317	IPB001271	Mammalian defensin	19.97 2.47e-09 242-270
1317	IPB001271	Mammalian defensin	19.97 2.80e-09 157-185
1317	IPB001271	Mammalian defensin	19.97 3.29e-10 56-84
1317	IPB001271	Mammalian defensin	19.97 4.76e-09 162-190
1317	IPB001271	Mammalian defensin	19.97 5.09e-09 26-54
1317	IPB001271	Mammalian defensin	19.97 5.09e-12 148-176
1317	IPB001271	Mammalian defensin	19.97 5.91e-09 66-94
1317	IPB001271	Mammalian defensin	19.97 6.25e-11 57-85
1317	IPB001271	Mammalian defensin	19.97 6.29e-10 51-79
1317	IPB001271	Mammalian defensin	19.97 6.47e-10 52-80
1317	IPB001271	Mammalian defensin	19.97 6.65e-10 158-186
1317	IPB001271	Mammalian defensin	19.97 7.05e-09 227-255
1317	IPB001271	Mammalian defensin	19.97 7.22e-09 167-195
1317	IPB001271	Mammalian defensin	19.97 7.53e-10 17-45

Table 3B  
406

SEQ ID NO:	Database entry ID	Description	Results*
1317	IPB001271	Mammalian defensin	19.97 7.71e-10 147-175
1317	IPB001271	Mammalian defensin	19.97 7.87e-09 202-230
1317	IPB001271	Mammalian defensin	19.97 8.20e-09 183-211
1317	IPB001271	Mammalian defensin	19.97 9.18e-09 192-220
1317	IPB001271	Mammalian defensin	19.97 9.35e-09 81-109
1317	IPB001271	Mammalian defensin	19.97 9.67e-09 123-151
1317	IPB001762A	Disintegrin	23.93 4.88e-09 91-131
1317	IPB001762A	Disintegrin	23.93 7.86e-12 99-139
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.00e-14 161-194
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.41e-18 201-234
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.66e-17 166-199
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.72e-10 232-265
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.88e-15 196-229
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.16e-16 50-83
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.35e-09 5-38
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.44e-12 80-113
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.44e-14 106-139
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.67e-16 131-164
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.67e-16 146-179
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.84e-10 251-284
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.12e-17 40-73
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.42e-09 97-130
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.42e-13 15-48
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.43e-19 171-204
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.43e-19 206-239
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.57e-16 25-58
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.81e-09 41-74
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.86e-19 226-259
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.88e-15 176-209
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.00e-09 81-114
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.23e-13 211-244
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.36e-14 45-78
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.44e-17 181-214
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.71e-17 30-63
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.87e-09 112-145
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.89e-12 111-144
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.00e-19 191-224
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.27e-13 96-129
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.35e-09 177-210
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.40e-10 26-59
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.50e-15 101-134
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.71e-11 151-184
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.73e-13 60-93
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.14e-16 35-68
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.28e-14 75-108
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.61e-09 237-270
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.66e-16 136-169
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.66e-16 65-98
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.66e-16 70-103
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.71e-09 36-69
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.73e-18 156-189
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.79e-11 221-254
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.00e-11 20-53

Table 3B  
407

SEQ ID NO:	Database entry ID	Description	Results*
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.11e-11 126-159
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.17e-16 121-154
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.22e-17 216-249
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.44e-12 186-219
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.58e-09 76-109
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.77e-09 182-215
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.97e-09 142-175
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.00e-12 236-269
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.07e-16 55-88
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.47e-10 246-279
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.67e-10 6-39
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.93e-11 141-174
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.13e-15 231-264
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.33e-12 241-274
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.81e-09 207-240
1317	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.90e-09 85-118
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.00e-09 160-174
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.45e-12 190-204
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.45e-12 225-239
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.63e-09 89-103
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 2.75e-10 125-139
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 3.11e-11 54-68
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 3.81e-11 120-134
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.64e-11 49-63
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.69e-09 180-194
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.69e-09 215-229
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.84e-10 19-33
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.92e-11 155-169
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.07e-09 69-83
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.58e-09 185-199
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.58e-09 220-234
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.96e-09 59-73
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 7.08e-09 150-164
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 7.08e-09 195-209
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 7.97e-09 79-93
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 8.52e-10 170-184
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 8.52e-10 205-219
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.25e-12 64-78
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.33e-10 130-144
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.37e-09 94-108
1317	IPB002494B	Keratin, high sulfur B2 protein	10.58 9.62e-09 240-254
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.12e-21 154-197
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.23e-18 13-56
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.35e-10 89-132
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.38e-12 22-65
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.59e-09 72-115
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.71e-20 224-267
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.73e-11 27-70
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.91e-11 37-80
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.11e-22 209-252
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.13e-12 84-127
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.22e-10 123-166
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.22e-12 168-211

Table 3B  
408

SEQ ID NO:	Database entry ID	Description	Results*
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.29e-29 184-227
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.31e-10 42-85
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.45e-15 99-142
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.48e-19 33-76
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.50e-14 234-277
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.55e-15 98-141
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.57e-29 58-101
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.60e-09 47-90
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.64e-11 143-186
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.00e-23 179-222
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.09e-11 133-176
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.36e-09 223-266
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.38e-23 204-247
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.44e-09 238-281
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.86e-09 218-261
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.90e-32 194-237
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.97e-10 12-55
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.06e-10 188-231
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.18e-11 178-221
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.36e-11 67-110
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.36e-21 199-242
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.49e-18 144-187
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.52e-28 43-86
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.55e-27 229-272
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.67e-10 158-201
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.84e-12 198-241
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.04e-09 103-146
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.10e-25 169-212
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.19e-10 153-196
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.50e-14 94-137
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.50e-18 83-126
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.50e-25 68-111
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.90e-14 244-287
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.16e-12 213-256
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.16e-21 214-257
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.19e-32 159-202
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.25e-23 174-217
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.35e-19 239-282
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.57e-16 78-121
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.74e-18 129-172
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.91e-12 62-105
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.96e-17 119-162
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.00e-11 173-216
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.00e-12 128-171
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.12e-10 208-251
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.19e-28 124-167
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.21e-16 3-46
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.30e-18 73-116
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.43e-24 134-177
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.50e-23 53-96
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.55e-10 32-75
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.64e-11 193-236
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.73e-10 148-191



Table 3B  
409

SEQ ID NO:	Database entry ID	Description	Results*
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.75e-20 114-157
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.88e-23 189-232
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.99e-20 109-152
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.03e-12 163-206
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.05e-31 149-192
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.15e-22 28-71
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.22e-12 8-51
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.25e-23 164-207
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.32e-21 23-66
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.81e-25 18-61
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.91e-09 77-120
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.92e-17 249-292
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.02e-17 88-131
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.04e-10 57-100
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.24e-09 19-62
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.32e-13 203-246
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.41e-20 63-106
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.43e-19 139-182
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.51e-22 38-81
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.73e-26 48-91
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.74e-10 108-151
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.82e-11 52-95
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.86e-28 219-262
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.89e-16 93-136
1317	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.90e-14 104-147
1317	PR00876B	Nematode metallothionein signature II	7.66 9.76e-09 113-126
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 1.00e-08 42-87
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 2.10e-09 125-170
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 2.74e-09 34-79
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 3.59e-12 124-169
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 4.21e-12 119-164
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 4.58e-09 32-77
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 4.67e-09 94-139
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.04e-09 88-133
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.13e-09 101-146
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.16e-10 100-145
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.20e-11 109-154
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.40e-11 120-165
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.65e-10 114-159
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.68e-09 110-155
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.78e-09 104-149
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 5.96e-09 55-100
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 6.03e-10 98-143
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 6.50e-11 35-80
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 6.59e-12 115-160
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 6.97e-09 111-156
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 7.61e-09 43-88
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 9.23e-10 83-128
1318	IPB000006	Vertebrate metallothionein, family 1	13.41 9.67e-13 93-138
1318	IPB000254	Cellulose-binding domain, fungal type	18.11 6.86e-09 94-124
1318	IPB000359B	Cystine-knot domain	19.26 3.25e-12 147-165
1318	IPB000359B	Cystine-knot domain	19.26 5.29e-13 127-145
1318	IPB000726A	Glycoside hydrolase family 19	14.41 9.88e-09 3-29

Table 3B  
410

SEQ ID NO:	Database entry ID	Description	Results*
1318	IPB000726B	Glycoside hydrolase family 19	5.39 3.40e-10 16-26
1318	IPB000726B	Glycoside hydrolase family 19	5.39 3.40e-10 9-19
1318	IPB001007C	von Willebrand factor, type C repeat	20.89 4.27e-09 111-132
1318	IPB001169K	Integrin beta, C-terminus	27.45 3.07e-09 118-160
1318	IPB001169K	Integrin beta, C-terminus	27.45 5.62e-09 25-67
1318	IPB001169K	Integrin beta, C-terminus	27.45 8.91e-09 35-77
1318	IPB001271	Mammalian defensin	19.97 5.91e-09 142-170
1318	IPB001271	Mammalian defensin	19.97 6.89e-09 107-135
1318	IPB001271	Mammalian defensin	19.97 8.41e-10 21-49
1318	IPB001762A	Disintegrin	23.93 3.03e-13 112-152
1318	IPB001762A	Disintegrin	23.93 4.18e-09 113-153
1318	IPB001762A	Disintegrin	23.93 5.60e-10 103-143
1318	IPB001762A	Disintegrin	23.93 5.60e-10 123-163
1318	IPB002174A	Furin-like cysteine rich region	30.51 1.97e-10 83-114
1318	IPB002174A	Furin-like cysteine rich region	30.51 3.92e-10 93-124
1318	IPB002174A	Furin-like cysteine rich region	30.51 8.20e-09 73-104
1318	IPB002174A	Furin-like cysteine rich region	30.51 8.88e-09 17-48
1318	IPB002174A	Furin-like cysteine rich region	30.51 9.27e-10 3-34
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.11e-11 101-134
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.13e-17 130-163
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.31e-10 116-149
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.78e-12 105-138
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.00e-12 25-58
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.16e-09 31-64
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.63e-15 140-173
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.32e-18 115-148
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.44e-16 135-168
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.21e-16 120-153
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.48e-10 100-133
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.04e-13 91-124
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.81e-10 30-63
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.32e-09 106-139
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.67e-12 14-47
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.89e-12 20-53
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.03e-10 96-129
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.65e-10 121-154
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.35e-09 21-54
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.88e-15 111-144
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.03e-09 26-59
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.22e-12 86-119
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.25e-11 24-57
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.57e-19 110-143
1318	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.73e-18 125-158
1318	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.51e-09 124-138
1318	IPB002494B	Keratin, high sulfur B2 protein	10.58 2.27e-09 159-173
1318	IPB002494B	Keratin, high sulfur B2 protein	10.58 3.55e-12 139-153
1318	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.70e-12 129-143
1318	IPB002494B	Keratin, high sulfur B2 protein	10.58 8.45e-11 149-163
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.27e-11 114-157
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.27e-11 99-142
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.41e-27 128-171
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.70e-10 104-147
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.08e-13 98-141

Table 3B  
411

SEQ ID NO:	Database entry ID	Description	Results*
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.48e-18 113-156
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.88e-10 107-150
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.09e-11 112-155
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.70e-09 102-145
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.84e-21 108-151
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.10e-19 123-166
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.55e-11 17-60
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.77e-17 103-146
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.82e-11 109-152
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.87e-13 23-66
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.81e-09 94-137
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.84e-12 122-165
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.40e-09 48-91
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.57e-29 118-161
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.69e-10 88-131
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.69e-12 117-160
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.33e-09 93-136
1318	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.71e-13 127-170
1318	IPB003571B	Snake toxin	18.08 4.72e-09 32-55
1318	PR00874C	Fungi-IV metallothionein signature III	4.37 9.74e-10 95-109
1318	PR00876B	Nematode metallothionein signature II	7.66 2.27e-10 107-120
1319	IPB000822	Zinc finger, C2H2 type	14.67 3.08e-18 175-200
1319	IPB000822	Zinc finger, C2H2 type	14.67 4.75e-19 147-172
1319	IPB000822	Zinc finger, C2H2 type	14.67 4.94e-14 119-144
1319	IPB000822	Zinc finger, C2H2 type	14.67 7.92e-18 63-88
1319	IPB000822	Zinc finger, C2H2 type	14.67 8.36e-21 203-228
1319	IPB000822	Zinc finger, C2H2 type	14.67 8.36e-21 91-116
1319	IPB000822	Zinc finger, C2H2 type	14.67 9.31e-18 35-60
1319	IPB001222	TFIIS zinc ribbon domain	24.63 6.70e-09 175-211
1319	IPB001275	DM DNA binding domain	19.17 3.43e-13 79-118
1319	IPB001275	DM DNA binding domain	19.17 4.77e-10 163-202
1319	IPB001275	DM DNA binding domain	19.17 6.91e-11 51-90
1319	IPB001275	DM DNA binding domain	19.17 7.20e-10 107-146
1319	IPB001275	DM DNA binding domain	19.17 9.09e-09 23-62
1319	IPB001275	DM DNA binding domain	19.17 9.10e-14 135-174
1319	IPB001275	DM DNA binding domain	19.17 9.29e-13 191-230
1319	IPB002867C	Cysteine-rich domain (C6HC)	19.46 8.34e-09 32-49
1319	PR00048A	C2H2-type zinc finger signature I	9.94 2.23e-10 200-213
1319	PR00048A	C2H2-type zinc finger signature I	9.94 3.84e-11 60-73
1319	PR00048A	C2H2-type zinc finger signature I	9.94 4.38e-09 88-101
1319	PR00048A	C2H2-type zinc finger signature I	9.94 5.50e-10 32-45
1319	PR00048A	C2H2-type zinc finger signature I	9.94 8.77e-10 172-185
1319	PR00048A	C2H2-type zinc finger signature I	9.94 9.18e-10 144-157
1319	PR00048B	C2H2-type zinc finger signature II	5.52 1.00e-08 132-141
1319	PR00048B	C2H2-type zinc finger signature II	5.52 2.50e-09 48-57
1319	PR00048B	C2H2-type zinc finger signature II	5.52 3.57e-10 160-169
1319	PR00048B	C2H2-type zinc finger signature II	5.52 4.00e-11 104-113
1319	PR00048B	C2H2-type zinc finger signature II	5.52 5.50e-09 76-85
1319	PR00048B	C2H2-type zinc finger signature II	5.52 6.14e-10 188-197
1319	PR00048B	C2H2-type zinc finger signature II	5.52 7.75e-11 216-225
1320	IPB002041A	GTP-binding nuclear protein Ran family	11.12 9.17e-10 4-48
1320	PR00449A	Transforming protein P21 ras signature I	12.48 1.29e-14 4-25
1321	IPB000006	Vertebrate metallothionein, family I	13.41 6.00e-11 71-116

Table 3B  
412

SEQ ID NO:	Database entry ID	Description	Results*
1321	IPB000006	Vertebrate metallothionein, family 1	13.41 9.72e-09 146-191
1321	IPB000726B	Glycoside hydrolase family 19	5.39 2.54e-09 66-76
1321	IPB000967E	Zinc finger NF-X1 type	21.88 1.24e-09 95-135
1321	IPB001169K	Integrin beta, C-terminus	27.45 5.38e-09 56-98
1321	IPB001762A	Disintegrin	23.93 9.29e-09 69-109
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.56e-12 161-194
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.06e-09 146-179
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 2.94e-09 103-136
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.00e-14 81-114
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 4.19e-09 72-105
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.26e-09 67-100
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.38e-13 156-189
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 5.56e-14 71-104
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.32e-09 166-199
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 6.56e-12 123-156
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.34e-10 124-157
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.39e-09 82-115
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.48e-09 151-184
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.58e-09 57-90
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.60e-14 118-151
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.97e-09 86-119
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.33e-12 113-146
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.35e-09 62-95
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.97e-16 76-109
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.13e-09 147-180
1321	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.14e-11 66-99
1321	IPB002494B	Keratin, high sulfur B2 protein	10.58 1.13e-09 70-84
1321	IPB002494B	Keratin, high sulfur B2 protein	10.58 3.53e-11 122-136
1321	IPB002494B	Keratin, high sulfur B2 protein	10.58 4.17e-09 142-156
1321	IPB002494B	Keratin, high sulfur B2 protein	10.58 4.17e-09 22-36
1321	IPB002494B	Keratin, high sulfur B2 protein	10.58 5.56e-09 137-151
1321	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.58e-09 80-94
1321	IPB002494B	Keratin, high sulfur B2 protein	10.58 6.78e-10 95-109
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.50e-14 21-64
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.59e-09 70-113
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 1.96e-10 6-49
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.01e-09 36-79
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.20e-21 74-117
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.27e-13 164-207
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.55e-11 26-69
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.70e-14 159-202
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.86e-13 149-192
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.10e-09 131-174
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.10e-09 75-118
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 3.30e-14 126-169
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.19e-12 84-127
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 4.36e-17 116-159
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.36e-11 144-187
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.61e-16 79-122
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 5.79e-09 145-188
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.06e-12 111-154
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.53e-12 11-54
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 6.91e-12 89-132

Table 3B  
413

SEQ ID NO:	Database entry ID	Description	Results*
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.10e-15 64-107
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.31e-15 121-164
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.82e-11 169-212
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.85e-13 59-102
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 7.99e-10 139-182
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 8.76e-18 154-197
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.07e-09 27-70
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.44e-12 31-74
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.66e-18 69-112
1321	IPB002494C	Keratin, high sulfur B2 protein	14.46 9.73e-11 106-149
1321	IPB002919A	Trypsin Inhibitor-like cysteine rich domain	15.56 6.57e-09 24-36
1327	IPB000118B	Granulin	7.94 1.30e-10 29-67
1327	IPB000118G	Granulin	12.18 8.85e-09 18-66
1327	IPB000203A	GPS domain	18.40 4.65e-09 44-74
1327	IPB000203A	GPS domain	18.40 5.57e-09 86-116
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 1.64e-09 58-103
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 2.70e-11 54-99
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 3.39e-09 70-115
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 3.66e-09 75-120
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 4.87e-10 51-96
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 5.59e-09 57-102
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 6.00e-11 50-95
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 6.42e-09 65-110
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 7.98e-09 40-85
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 8.07e-09 55-100
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 8.48e-13 47-92
1329	IPB000006	Vertebrate metallothionein, family 1	13.41 8.94e-10 74-119
1329	IPB000254	Cellulose-binding domain, fungal type	18.11 7.00e-09 43-73
1329	IPB000564A	2Fe-2S Ferredoxin	17.31 7.80e-09 1-19
1329	IPB000817A	Prion protein	8.34 1.85e-09 13-55
1329	IPB000817A	Prion protein	8.34 4.79e-09 9-51
1329	IPB000817A	Prion protein	8.34 6.87e-09 12-54
1329	IPB000817A	Prion protein	8.34 8.64e-11 6-48
1329	IPB000817A	Prion protein	8.34 9.01e-10 10-52
1329	IPB000967D	Zinc finger NF-X1 type	10.42 7.77e-09 75-110
1329	IPB001169K	Integrin beta, C-terminus	27.45 3.19e-09 57-99
1329	IPB001169K	Integrin beta, C-terminus	27.45 8.05e-09 72-114
1329	IPB001169K	Integrin beta, C-terminus	27.45 9.15e-09 16-58
1329	IPB001271	Mammalian defensin	19.97 3.29e-10 66-94
1329	IPB001271	Mammalian defensin	19.97 4.44e-09 73-101
1329	IPB001271	Mammalian defensin	19.97 5.42e-09 82-110
1329	IPB001271	Mammalian defensin	19.97 7.35e-10 44-72
1329	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.10e-09 4-56
1329	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.80e-09 1-53
1329	IPB002174A	Furin-like cysteine rich region	30.51 2.58e-09 51-82
1329	IPB002174A	Furin-like cysteine rich region	30.51 4.38e-09 34-65
1329	IPB002174A	Furin-like cysteine rich region	30.51 4.97e-11 2-33
1329	IPB002174A	Furin-like cysteine rich region	30.51 5.28e-09 6-37
1329	IPB002174A	Furin-like cysteine rich region	30.51 6.35e-10 26-57
1329	IPB002174A	Furin-like cysteine rich region	30.51 7.62e-11 30-61
1329	IPB002174A	Furin-like cysteine rich region	30.51 9.27e-10 18-49

Table 3B

414

SEQ ID NO:	Database entry ID	Description	Results*
1329	IPB002174A	Furin-like cysteine rich region	30.51 9.27e-10 22-53
1329	IPB002221B	WAP-type (Whey Acidic Protein) four-disulfide core domain	17.12 1.00e-08 63-84
1329	IPB002494A	Keratin, high sulfur B2 protein	12.44 1.39e-09 44-77
1329	IPB002494A	Keratin, high sulfur B2 protein	12.44 3.23e-09 34-67
1329	IPB002494A	Keratin, high sulfur B2 protein	12.44 7.75e-10 85-118
1329	IPB002494A	Keratin, high sulfur B2 protein	12.44 8.67e-10 47-80
1329	IPB002494A	Keratin, high sulfur B2 protein	12.44 9.03e-09 40-73
1329	IPB002494C	Keratin, high sulfur B2 protein	14.46 2.31e-10 60-103
1329	IPB002551P	Coronavirus S1 glycoprotein	13.91 7.68e-09 79-120
1329	IPB002867D	Cysteine-rich domain (C6HC)	24.88 6.49e-09 53-84
1329	IPB002919A	Trypsin Inhibitor-like cysteine rich domain	15.56 3.57e-09 67-79
1329	PR00858B	Crustacean metallothionein signature II	5.93 1.48e-09 55-73
1329	PR00872A	Diptera (Drosophila) metallothionein signature I	5.17 7.41e-09 99-111
1329	PR01117A	CLC-6 chloride channel signature I	7.79 9.47e-09 66-78
1329	PR01131B	Connexin36 (Cx36) signature II	3.45 5.12e-09 42-54
1329	PR01228C	Eggshell protein signature III	5.69 1.22e-10 17-32
1329	PR01228C	Eggshell protein signature III	5.69 1.22e-10 21-36
1329	PR01228C	Eggshell protein signature III	5.69 1.22e-10 25-40
1329	PR01228C	Eggshell protein signature III	5.69 1.22e-10 29-44
1329	PR01228C	Eggshell protein signature III	5.69 1.22e-10 33-48
1329	PR01228C	Eggshell protein signature III	5.69 1.22e-10 37-52
1329	PR01228C	Eggshell protein signature III	5.69 2.34e-09 41-56
1329	PR01228C	Eggshell protein signature III	5.69 5.86e-09 18-33
1329	PR01228C	Eggshell protein signature III	5.69 5.86e-09 22-37
1329	PR01228C	Eggshell protein signature III	5.69 5.86e-09 26-41
1329	PR01228C	Eggshell protein signature III	5.69 5.86e-09 30-45
1329	PR01228C	Eggshell protein signature III	5.69 5.86e-09 34-49
1329	PR01228C	Eggshell protein signature III	5.69 5.86e-09 38-53
1329	PR01228C	Eggshell protein signature III	5.69 6.48e-09 13-28
1330	IPB000221	Protamine P1	5.48 1.00e-08 76-102
1330	IPB000221	Protamine P1	5.48 1.00e-09 83-109
1330	IPB000221	Protamine P1	5.48 1.46e-09 65-91
1330	IPB000221	Protamine P1	5.48 2.19e-11 103-129
1330	IPB000221	Protamine P1	5.48 2.59e-11 64-90
1330	IPB000221	Protamine P1	5.48 2.73e-10 118-144
1330	IPB000221	Protamine P1	5.48 2.97e-12 74-100
1330	IPB000221	Protamine P1	5.48 3.31e-09 109-135
1330	IPB000221	Protamine P1	5.48 3.31e-09 122-148
1330	IPB000221	Protamine P1	5.48 3.91e-11 78-104
1330	IPB000221	Protamine P1	5.48 4.70e-10 62-88
1330	IPB000221	Protamine P1	5.48 4.70e-10 94-120
1330	IPB000221	Protamine P1	5.48 5.15e-09 107-133
1330	IPB000221	Protamine P1	5.48 5.27e-09 52-78
1330	IPB000221	Protamine P1	5.48 6.16e-11 92-118
1330	IPB000221	Protamine P1	5.48 6.19e-09 116-142
1330	IPB000221	Protamine P1	5.48 6.43e-11 99-125
1330	IPB000221	Protamine P1	5.48 6.88e-09 111-137
1330	IPB000221	Protamine P1	5.48 6.88e-09 97-123
1330	IPB000221	Protamine P1	5.48 7.62e-11 60-86
1330	IPB000221	Protamine P1	5.48 8.89e-10 101-127
1330	IPB000221	Protamine P1	5.48 9.30e-12 63-89

Table 3B  
415

SEQ ID NO:	Database entry ID	Description	Results*
1330	IPB000271	Ribosomal protein L34	15.87 9.78e-09 111-148
1330	IPB000492B	Protamine 2 (PRM2)	5.26 3.84e-09 75-109
1330	IPB000492B	Protamine 2 (PRM2)	5.26 5.88e-11 98-132
1330	IPB000492B	Protamine 2 (PRM2)	5.26 6.38e-09 94-128
1330	IPB000492B	Protamine 2 (PRM2)	5.26 6.67e-09 107-141
1330	IPB000492B	Protamine 2 (PRM2)	5.26 6.97e-10 103-137
1330	IPB000492B	Protamine 2 (PRM2)	5.26 7.75e-09 77-111
1330	IPB000492B	Protamine 2 (PRM2)	5.26 8.12e-10 106-140
1330	IPB000492B	Protamine 2 (PRM2)	5.26 8.34e-09 65-99
1330	IPB000492B	Protamine 2 (PRM2)	5.26 8.53e-10 105-139
1330	IPB000492B	Protamine 2 (PRM2)	5.26 9.06e-10 78-112
1330	IPB000492B	Protamine 2 (PRM2)	5.26 9.35e-11 102-136
1330	IPB000492B	Protamine 2 (PRM2)	5.26 9.35e-11 79-113
1330	IPB000492B	Protamine 2 (PRM2)	5.26 9.69e-10 100-134
1330	IPB000492B	Protamine 2 (PRM2)	5.26 9.90e-09 111-145
1330	IPB003134F	Repeat in HS1/Cortactin	15.66 7.21e-09 183-231
1330	PR00055C	HIV TAT domain signature III	9.12 5.92e-09 16-32
1331	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 2.80e-10 54-91
1332	IPB002223	Pancreatic trypsin inhibitor (Kunitz) family	17.66 2.61e-23 92-126
1332	PR00759A	Basic protease (Kunitz-type) inhibitor family signature I	14.67 5.71e-10 73-87
1332	PR00759B	Basic protease (Kunitz-type) inhibitor family signature II	12.35 1.28e-10 101-111
1332	PR00759C	Basic protease (Kunitz-type) inhibitor family signature III	12.43 2.24e-11 111-126
1333	IPB000822	Zinc finger, C2H2 type	14.67 1.75e-19 201-226
1333	IPB000822	Zinc finger, C2H2 type	14.67 1.75e-20 341-366
1333	IPB000822	Zinc finger, C2H2 type	14.67 2.13e-15 369-394
1333	IPB000822	Zinc finger, C2H2 type	14.67 5.50e-20 229-254
1333	IPB000822	Zinc finger, C2H2 type	14.67 5.91e-21 285-310
1333	IPB000822	Zinc finger, C2H2 type	14.67 7.00e-20 313-338
1333	IPB000822	Zinc finger, C2H2 type	14.67 9.10e-22 257-282
1333	IPB001275	DM DNA binding domain	19.17 1.29e-13 245-284
1333	IPB001275	DM DNA binding domain	19.17 2.34e-12 329-368
1333	IPB001275	DM DNA binding domain	19.17 2.95e-10 217-256
1333	IPB001275	DM DNA binding domain	19.17 3.51e-09 357-396
1333	IPB001275	DM DNA binding domain	19.17 7.00e-13 273-312
1333	IPB001275	DM DNA binding domain	19.17 7.15e-14 301-340
1333	IPB001909	KRAB box	17.37 9.10e-31 8-42
1333	PR00048A	C2H2-type zinc finger signature I	9.94 1.00e-13 198-211
1333	PR00048A	C2H2-type zinc finger signature I	9.94 1.64e-13 254-267
1333	PR00048A	C2H2-type zinc finger signature I	9.94 1.64e-13 282-295
1333	PR00048A	C2H2-type zinc finger signature I	9.94 6.21e-11 226-239
1333	PR00048A	C2H2-type zinc finger signature I	9.94 6.79e-13 366-379
1333	PR00048A	C2H2-type zinc finger signature I	9.94 9.36e-13 310-323
1333	PR00048A	C2H2-type zinc finger signature I	9.94 9.53e-11 338-351
1333	PR00048B	C2H2-type zinc finger signature II	5.52 1.00e-09 270-279
1333	PR00048B	C2H2-type zinc finger signature II	5.52 2.29e-10 326-335
1333	PR00048B	C2H2-type zinc finger signature II	5.52 2.50e-11 242-251
1333	PR00048B	C2H2-type zinc finger signature II	5.52 7.00e-09 298-307
1333	PR00048B	C2H2-type zinc finger signature II	5.52 7.00e-12 354-363
1333	PR00048B	C2H2-type zinc finger signature II	5.52 7.50e-09 214-223

Table 3B

416

SEQ ID NO:	Database entry ID	Description	Results*
1336	PR00806B	Vinculin signature II	4.27 9.04e-09 333-346
1336	PR00806B	Vinculin signature II	4.27 9.64e-09 332-345
1336	PR01217G	Proline rich extensin signature VII	4.02 7.47e-09 399-424
1337	PR00806B	Vinculin signature II	4.27 9.04e-09 333-346
1337	PR00806B	Vinculin signature II	4.27 9.64e-09 332-345
1337	PR01217G	Proline rich extensin signature VII	4.02 7.47e-09 399-424
1338	IPB001526A	Ly-6/u-PAR domain	13.24 8.33e-09 7-22
1338	PR01226A	Expansin signature I	5.84 6.11e-09 259-273
1339	IPB000729D	PMP-22/EMP/MP20 family	18.96 7.30e-10 182-209
1340	IPB000729D	PMP-22/EMP/MP20 family	18.96 7.30e-10 173-200
1341	IPB003397B	Mitochondrial import inner membrane translocase subunit Tim17	21.18 9.79e-10 60-105
1343	IPB000353B	Class II histocompatibility antigen, beta chain, beta-1 domain	19.16 4.94e-16 187-236
1343	IPB001003B	MHC Class II, alpha chain, alpha-1 domain	14.72 9.90e-10 195-238
1343	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 8.50e-16 208-230
1344	IPB002499N	Major Vault Protein repeat	12.73 9.47e-09 220-273
1345	PR00462B	Fungal lignin peroxidase family signature II	14.83 4.74e-09 34-52
1347	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 3.77e-09 202-239
1347	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 8.83e-11 306-343
1347	PR01236A	Tumour necrosis factor beta (lymphotoxin-alpha) signature I	4.92 5.60e-09 3-19
1348	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 1.36e-10 403-440
1350	PR01218B	Pistil-specific extensin-like signature II	8.47 4.88e-09 115-138
1353	PR00019A	Leucine-rich repeat signature I	11.72 7.33e-10 193-206
1356	IPB002896F	Herpesvirus glycoprotein D	21.08 9.22e-09 327-362
1356	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 5.15e-09 276-313
1357	IPB002159A	CD36 family	36.08 5.26e-39 258-311
1357	IPB002159B	CD36 family	29.25 5.13e-34 334-379
1359	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 1.00e-10 160-182
1359	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 6.09e-11 206-243
1363	PR01432K	Rabaptin signature XI	2.19 8.43e-09 873-895
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 1.00e-15 200-237
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 1.20e-10 167-204
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 1.25e-16 191-228
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.00e-13 203-240
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.20e-20 179-216
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.80e-18 176-213
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 4.29e-19 182-219
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 6.58e-10 209-246
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 7.15e-09 164-201
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 7.85e-16 173-210
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 8.20e-20 185-222
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 8.99e-11 206-243
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.03e-15 197-234
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.11e-13 170-207
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.37e-16 194-231



Table 3B

417

SEQ ID NO:	Database entry ID	Description	Results*
1366	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.71e-19 188-225
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.12e-19 174-227
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 2.73e-20 186-239
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 3.59e-17 171-224
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 4.58e-16 189-242
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 5.11e-10 153-206
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 5.31e-12 159-212
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 5.77e-16 192-245
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 5.98e-15 162-215
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.06e-14 165-218
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.18e-20 177-230
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.46e-14 195-248
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.66e-10 201-254
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.87e-13 198-251
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.36e-19 168-221
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.55e-13 156-209
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 8.78e-12 150-203
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.16e-24 180-233
1366	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.60e-22 183-236
1366	IPB001073A	Complement C1q protein	22.14 1.00e-13 202-236
1366	IPB001073A	Complement C1q protein	22.14 1.46e-16 193-227
1366	IPB001073A	Complement C1q protein	22.14 2.76e-15 196-230
1366	IPB001073A	Complement C1q protein	22.14 3.77e-16 190-224
1366	IPB001073A	Complement C1q protein	22.14 3.90e-10 208-242
1366	IPB001073A	Complement C1q protein	22.14 4.69e-16 175-209
1366	IPB001073A	Complement C1q protein	22.14 5.83e-15 178-212
1366	IPB001073A	Complement C1q protein	22.14 6.18e-19 184-218
1366	IPB001073A	Complement C1q protein	22.14 7.51e-13 181-215
1366	IPB001073A	Complement C1q protein	22.14 8.02e-15 205-239
1366	IPB001073A	Complement C1q protein	22.14 8.40e-14 199-233
1366	IPB001073A	Complement C1q protein	22.14 8.78e-10 172-206
1366	IPB001073A	Complement C1q protein	22.14 8.85e-13 187-221
1366	IPB001073B	Complement C1q protein	20.88 7.26e-29 281-315
1366	IPB001073C	Complement C1q protein	13.07 1.87e-14 349-368
1366	IPB001073D	Complement C1q protein	7.60 8.20e-13 383-392
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 1.44e-18 169-221
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 1.52e-23 181-233
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.08e-14 160-212
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.00e-18 166-218
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.20e-16 202-254
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.25e-12 157-209
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.72e-13 196-248
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.93e-17 193-245
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.97e-12 199-251
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.08e-21 175-227

Table 3B  
418

SEQ ID NO:	Database entry ID	Description	Results*
		4 procollagen	
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.45e-16 190-242
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.74e-19 184-236
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.01e-22 172-224
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.24e-09 154-206
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.40e-22 178-230
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.54e-11 148-200
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.86e-17 187-239
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.97e-16 163-215
1366	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 8.47e-12 151-203
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 1.57e-11 209-229
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 2.34e-09 182-202
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 2.82e-09 188-208
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 3.55e-09 179-199
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 4.00e-11 206-226
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 4.47e-10 191-211
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 5.01e-09 203-223
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 5.24e-10 194-214
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 7.30e-10 200-220
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 7.57e-11 197-217
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 7.93e-09 176-196
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 8.42e-09 185-205
1366	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 8.66e-09 212-232
1366	PR00007A	Complement C1Q domain signature I	20.64 6.54e-20 274-300
1366	PR00007B	Complement C1Q domain signature II	15.63 3.13e-14 301-320
1366	PR00007C	Complement C1Q domain signature III	16.13 2.62e-15 349-370
1366	PR00007D	Complement C1Q domain signature IV	9.66 1.95e-12 381-391
1367	IPB001427A	Pancreatic ribonuclease family	12.97 1.41e-10 31-40
1367	IPB001427B	Pancreatic ribonuclease family	31.44 4.30e-29 45-95
1367	IPB001427C	Pancreatic ribonuclease family	24.33 8.07e-20 83-113
1367	PR00794A	Pancreatic ribonuclease family signature I	15.68 1.00e-18 46-65
1367	PR00794B	Pancreatic ribonuclease family signature II	17.91 7.55e-16 66-85

Table 3B  
419

SEQ ID NO:	Database entry ID	Description	Results*
1367	PR00794D	Pancreatic ribonuclease family signature IV	19.74 6.40e-20 94-116
1368	IPB000171A	Bacterial-type phytoene dehydrogenase	19.28 9.72e-11 57-87
1368	IPB000927A	D-amino acid oxidase	8.56 7.91e-09 55-67
1368	IPB002135A	3-hydroxyacyl-CoA dehydrogenase	20.39 2.50e-10 58-83
1369	IPB000171A	Bacterial-type phytoene dehydrogenase	19.28 9.72e-11 57-87
1369	IPB000927A	D-amino acid oxidase	8.56 7.91e-09 55-67
1369	IPB002135A	3-hydroxyacyl-CoA dehydrogenase	20.39 2.50e-10 58-83
1371	PR00762A	Chloride channel signature I	14.76 7.95e-19 108-125
1371	PR00762B	Chloride channel signature II	11.84 2.97e-16 139-158
1371	PR00762C	Chloride channel signature III	9.42 3.57e-20 203-222
1371	PR00762D	Chloride channel signature IV	11.34 4.54e-19 440-460
1371	PR00762E	Chloride channel signature V	11.24 8.45e-16 475-491
1371	PR00762F	Chloride channel signature VI	15.29 4.18e-15 493-512
1371	PR00762G	Chloride channel signature VII	13.53 7.19e-12 531-545
1371	PR01117A	CLC-6 chloride channel signature I	7.79 1.09e-19 2-14
1371	PR01117C	CLC-6 chloride channel signature III	9.38 1.19e-17 54-66
1371	PR01117D	CLC-6 chloride channel signature IV	6.15 1.10e-13 159-168
1371	PR01117E	CLC-6 chloride channel signature V	7.29 1.10e-13 170-179
1372	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 1.69e-13 159-170
1372	PR00245A	Olfactory receptor signature I	10.98 2.50e-13 133-144
1372	PR00245B	Olfactory receptor signature II	13.73 5.70e-12 170-182
1372	PR00534A	Melanocortin receptor family signature I	12.77 8.04e-09 92-104
1372	PR00642C	EDG-1 sphingosine 1-phosphate receptor signature III	13.87 4.43e-09 89-101
1374	IPB000109A	PTR peptide transporters (PTR2)	10.85 3.79e-15 44-62
1374	IPB000109B	PTR peptide transporters (PTR2)	29.23 4.18e-23 67-119
1374	PR00308C	Type I antifreeze protein signature III	2.79 2.78e-09 20-29
1374	PR01471B	Histamine H3 receptor signature II	12.38 9.63e-09 24-42
1375	IPB000169A	Eukaryotic thiol (cysteine) proteases active site	10.17 1.00e-14 132-141
1375	IPB000169C	Eukaryotic thiol (cysteine) proteases active site	8.11 2.93e-11 274-284
1375	IPB000169D	Eukaryotic thiol (cysteine) proteases active site	8.56 9.44e-19 295-311
1375	PR00704C	Calpain cysteine protease (C2) family signature III	11.35 6.01e-09 132-148
1375	PR00705A	Papain cysteine protease (C1) family signature I	11.22 2.80e-21 132-147
1375	PR00705B	Papain cysteine protease (C1) family signature II	10.02 1.45e-10 276-286
1375	PR00705C	Papain cysteine protease (C1) family signature III	13.28 1.00e-09 295-301
1376	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 9.05e-10 125-136
1376	IPB000276B	Rhodopsin-like GPCR superfamily	4.97 5.74e-09 201-212
1376	PR00237B	Rhodopsin-like GPCR superfamily signature II	12.45 7.87e-09 66-87
1376	PR00237C	Rhodopsin-like GPCR superfamily signature III	14.77 8.50e-12 111-133
1376	PR00237E	Rhodopsin-like GPCR superfamily signature V	13.03 4.21e-11 193-216
1376	PR00237F	Rhodopsin-like GPCR superfamily signature VI	14.34 4.55e-11 397-421
1376	PR00237G	Rhodopsin-like GPCR superfamily signature VII	19.23 6.09e-09 436-462

Table 3B  
420

SEQ ID NO:	Database entry ID	Description	Results*
1376	PR01103B	Adrenergic receptor signature II	13.72 2.50e-10 139-147
1378	IPB000203A	GPS domain	18.40 9.74e-09 209-239
1378	IPB000203B	GPS domain	13.98 3.25e-15 325-346
1378	IPB000832C	G-protein coupled receptors family 2 (secretin-like)	19.53 1.00e-10 325-354
1378	PR00249C	Secretin-like GPCR superfamily signature III	15.44 3.32e-09 327-350
1378	PR01380A	Claudin-5 signature I	5.16 8.79e-09 265-273
1380	IPB002190A	MAGE family	31.63 4.40e-18 1-37
1380	IPB002190B	MAGE family	27.39 1.61e-26 109-153
1380	IPB002190C	MAGE family	33.01 9.12e-31 179-233
1380	IPB002190D	MAGE family	18.35 1.00e-40 239-288
1381	PR00169G	Potassium channel signature VII	11.30 6.38e-09 265-287
1381	PR01320B	Inward rectifier K <sup>+</sup> channel superfamily signature II	12.73 7.46e-11 162-184
1381	PR01333A	Two pore domain K <sup>+</sup> channel signature I	18.74 8.11e-28 168-196
1381	PR01333B	Two pore domain K <sup>+</sup> channel signature II	10.39 2.58e-09 171-180
1381	PR01333B	Two pore domain K <sup>+</sup> channel signature II	10.39 7.10e-10 280-289
1381	PR01463F	EAG/ELK/ERG potassium channel family signature VI	4.09 3.42e-09 268-285
1381	PR01499A	Trek K <sup>+</sup> channel signature I	8.28 7.99e-12 21-34
1381	PR01499B	Trek K <sup>+</sup> channel signature II	9.88 8.15e-21 68-85
1381	PR01499C	Trek K <sup>+</sup> channel signature III	9.14 7.07e-14 105-117
1381	PR01499D	Trek K <sup>+</sup> channel signature IV	8.87 9.05e-20 119-137
1381	PR01499E	Trek K <sup>+</sup> channel signature V	14.43 2.73e-14 146-158
1381	PR01499F	Trek K <sup>+</sup> channel signature VI	10.57 1.00e-24 196-215
1381	PR01499G	Trek K <sup>+</sup> channel signature VII	12.38 7.80e-22 216-235
1381	PR01499H	Trek K <sup>+</sup> channel signature VIII	9.00 6.71e-14 237-248
1381	PR01499I	Trek K <sup>+</sup> channel signature IX	11.15 4.73e-20 293-308
1381	PR01499J	Trek K <sup>+</sup> channel signature X	16.62 1.11e-16 315-326
1381	PR01499K	Trek K <sup>+</sup> channel signature XI	11.81 2.96e-20 328-344
1381	PR01499L	Trek K <sup>+</sup> channel signature XII	13.04 4.80e-17 364-378
1382	IPB000920C	Myelin P0 protein	15.78 6.50e-09 245-297
1382	IPB003784	BioY	27.20 1.22e-09 263-313
1382	PR00213E	Myelin P0 protein signature V	5.51 8.97e-09 263-287
1382	PR01228C	Eggshell protein signature III	5.69 5.14e-09 306-321
1385	IPB001039A	Major histocompatibility complex protein, Class I	17.17 8.89e-11 19-72
1388	IPB002130A	Cyclophilin-type peptidyl-prolyl cis-trans isomerase	13.87 1.53e-18 58-79
1388	IPB002130B	Cyclophilin-type peptidyl-prolyl cis-trans isomerase	21.20 2.93e-37 88-126
1388	IPB002130C	Cyclophilin-type peptidyl-prolyl cis-trans isomerase	16.92 9.50e-34 138-175
1388	PR00153A	Cyclophilin peptidyl-prolyl cis-trans isomerase signature I	14.60 1.00e-15 64-79
1388	PR00153B	Cyclophilin peptidyl-prolyl cis-trans isomerase signature II	12.94 5.24e-17 93-105
1388	PR00153C	Cyclophilin peptidyl-prolyl cis-trans isomerase signature III	10.79 1.00e-17 136-151
1388	PR00153D	Cyclophilin peptidyl-prolyl cis-trans isomerase signature IV	10.73 7.75e-16 151-163
1388	PR00153E	Cyclophilin peptidyl-prolyl cis-trans isomerase signature V	8.39 5.15e-10 164-179

Table 3B

421

SEQ ID NO:	Database entry ID	Description	Results*
1390	IPB000276B	Rhodopsin-like GPCR superfamily	4.97 6.06e-10 102-113
1390	IPB000276C	Rhodopsin-like GPCR superfamily	8.03 8.50e-11 162-173
1390	PR00237D	Rhodopsin-like GPCR superfamily signature IV	9.76 4.38e-09 58-79
1390	PR00237E	Rhodopsin-like GPCR superfamily signature V	13.03 6.50e-15 94-117
1390	PR00237F	Rhodopsin-like GPCR superfamily signature VI	14.34 5.91e-15 152-176
1390	PR00237G	Rhodopsin-like GPCR superfamily signature VII	19.23 5.70e-09 192-218
1390	PR00529D	Gonadotrophin releasing hormone receptor signature IV	22.60 7.38e-13 2-17
1390	PR00529F	Gonadotrophin releasing hormone receptor signature VI	5.71 5.09e-10 75-89
1390	PR00529H	Gonadotrophin releasing hormone receptor signature VIII	9.64 9.43e-11 176-196
1392	IPB003662A	General substrate transporters	18.97 8.76e-13 196-228
1393	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 1.69e-13 157-168
1393	PR00237C	Rhodopsin-like GPCR superfamily signature III	14.77 3.84e-10 143-165
1393	PR00245B	Olfactory receptor signature II	13.73 6.54e-11 168-180
1393	PR00245C	Olfactory receptor signature III	14.65 3.75e-11 215-231
1393	PR00245E	Olfactory receptor signature V	8.96 7.91e-10 322-333
1397	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 6.88e-09 294-331
1397	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 8.92e-10 487-524
1397	IPB003529B	Long hematopoietin receptor, gp130 family	20.65 2.31e-09 646-668
1397	PR00014C	Fibronectin type III repeat signature III	14.47 8.71e-09 585-603
1397	PR00014D	Fibronectin type III repeat signature IV	15.12 2.96e-09 1023-1037
1397	PR00014D	Fibronectin type III repeat signature IV	15.12 9.22e-09 701-715
1402	IPB000195B	RabGAP/TBC domain	14.35 5.15e-09 38-50
1402	PR00250G	Fungal pheromone mating factor STE2 GPCR signature VII	9.58 9.75e-09 21-42
1409	IPB001359H	Synapsin	22.58 9.63e-10 53-103
1409	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 1.43e-13 297-334
1409	IPB003531C	Short hematopoietin receptor family 1	15.87 9.38e-11 543-560
1413	PR01131B	Connexin36 (Cx36) signature II	3.45 2.41e-09 500-512
1414	IPB001993A	Mitochondrial energy transfer proteins (carrier protein)	16.91 7.00e-11 143-167
1414	IPB001993A	Mitochondrial energy transfer proteins (carrier protein)	16.91 7.50e-18 46-70
1414	IPB001993B	Mitochondrial energy transfer proteins (carrier protein)	9.00 5.15e-14 182-195
1414	PR00927E	Adenine nucleotide translocator 1 signature V	14.99 6.33e-09 74-95
1416	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 1.00e-09 6-43
1416	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 7.92e-09 88-125
1417	IPB000315A	B-box zinc finger superfamily	24.51 3.25e-10 30-53
1417	IPB001841	RING finger	10.69 1.56e-10 42-51
1418	IPB001285A	Synaptophysin/synaptoporin	12.19 1.00e-28 41-87

Table 3B  
422

SEQ ID NO:	Database entry ID	Description	Results*
1418	IPB001285C	Synaptophysin/synaptoporin	15.24 9.89e-23 125-165
1418	IPB001285D	Synaptophysin/synaptoporin	12.74 4.03e-13 166-200
1418	IPB001285E	Synaptophysin/synaptoporin	9.55 2.98e-24 201-247
1418	PR00220A	Synaptophysin/synaptoporin family signature I	13.69 7.60e-26 38-60
1418	PR00220B	Synaptophysin/synaptoporin family signature II	17.03 5.35e-23 62-87
1418	PR00220C	Synaptophysin/synaptoporin family signature III	11.86 2.85e-25 117-141
1418	PR00220D	Synaptophysin/synaptoporin family signature IV	10.60 6.84e-25 149-172
1418	PR00220E	Synaptophysin/synaptoporin family signature V	4.56 4.60e-23 216-234
1418	PR00489C	Frizzled protein signature III	8.19 2.36e-09 216-240
1419	IPB000729B	PMP-22/EMP/MP20 family	13.56 4.86e-11 49-59
1419	IPB000729C	PMP-22/EMP/MP20 family	37.83 9.01e-15 80-132
1419	IPB000729C	PMP-22/EMP/MP20 family	37.83 9.72e-10 87-139
1421	IPB001905A	Ammonium transporter family	22.03 8.47e-14 69-97
1421	IPB001905B	Ammonium transporter family	19.59 6.08e-17 127-159
1421	IPB001905C	Ammonium transporter family	15.52 5.35e-18 178-203
1421	IPB001905D	Ammonium transporter family	11.89 3.61e-19 217-240
1421	IPB001905F	Ammonium transporter family	17.91 1.00e-18 287-313
1421	PR00342B	Rhesus blood group protein signature II	14.04 5.57e-13 61-78
1421	PR00342C	Rhesus blood group protein signature III	11.52 5.03e-11 90-107
1421	PR00342D	Rhesus blood group protein signature IV	11.18 8.69e-13 128-144
1421	PR00342E	Rhesus blood group protein signature V	17.96 3.74e-10 151-174
1421	PR00342F	Rhesus blood group protein signature VI	8.73 3.88e-10 185-200
1421	PR00342G	Rhesus blood group protein signature VII	10.10 2.70e-20 220-238
1421	PR00342H	Rhesus blood group protein signature VIII	8.86 1.76e-12 250-272
1423	IPB001427A	Pancreatic ribonuclease family	12.97 5.09e-10 36-45
1423	IPB001427B	Pancreatic ribonuclease family	31.44 6.04e-34 50-100
1423	IPB001427C	Pancreatic ribonuclease family	24.33 2.17e-13 106-136
1423	PR00794A	Pancreatic ribonuclease family signature I	15.68 6.31e-13 51-70
1423	PR00794B	Pancreatic ribonuclease family signature II	17.91 6.34e-13 71-90
1423	PR00794C	Pancreatic ribonuclease family signature III	16.87 8.07e-14 96-114
1423	PR00794D	Pancreatic ribonuclease family signature IV	19.74 2.37e-09 117-139
1426	IPB000038A	Cell division/GTP binding protein	16.77 9.89e-09 447-488
1426	IPB001984C	ATP-dependent serine proteases, Lon family	12.76 3.32e-09 158-194
1426	PR00318A	Alpha G protein (transducin) signature I	6.88 5.19e-10 469-484
1426	PR00830A	Endopeptidase La (Lon) serine protease (S16) signature I	8.52 2.14e-11 165-184
1426	PR01100A	Shikimate kinase family signature I	8.30 1.00e-08 470-485
1427	IPB000483	Leucine rich repeat C-terminal domain	11.18 6.85e-13 164-178
1427	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 8.20e-10 259-296
1427	PR00019A	Leucine-rich repeat signature I	11.72 3.18e-09 52-65
1427	PR00019A	Leucine-rich repeat signature I	11.72 5.64e-09 77-90
1427	PR00019B	Leucine-rich repeat signature II	11.42 1.27e-09 25-38
1427	PR00019B	Leucine-rich repeat signature II	11.42 4.82e-09 49-62
1427	PR00019B	Leucine-rich repeat signature II	11.42 6.32e-11 74-87
1430	PR00462B	Fungal lignin peroxidase family signature II	14.83 4.74e-09 34-52
1435	IPB002610A	Rhomboid family	5.73 9.50e-10 185-194
1437	IPB001627A	Sema domain	16.97 9.34e-09 138-153

Table 3B  
423

SEQ ID NO:	Database entry ID	Description	Results*
1437	IPB001627C	Sema domain	21.13 7.11e-09 201-232
1437	IPB001627D	Sema domain	16.04 9.47e-17 249-264
1437	IPB001627F	Sema domain	22.05 8.50e-27 293-326
1437	IPB001627G	Sema domain	21.49 5.36e-22 346-379
1437	IPB001627J	Sema domain	11.43 2.97e-14 462-478
1437	IPB001627K	Sema domain	13.76 4.00e-14 540-552
1437	IPB002165A	Plexin repeat	15.07 3.73e-12 138-153
1437	IPB002165B	Plexin repeat	13.59 3.43e-12 175-184
1437	IPB002165C	Plexin repeat	18.49 6.32e-34 293-325
1437	IPB002165D	Plexin repeat	14.72 2.85e-15 540-552
1437	IPB002861B	Reeler domain	10.50 3.56e-09 890-918
1437	IPB002861B	Reeler domain	10.50 9.76e-12 644-672
1437	PR01303D	Plasmodium circumsporozoite protein signature IV	10.57 2.50e-09 708-725
1438	PR00170A	Voltage-gated Na <sup>+</sup> channel alpha subunit signature I	8.35 3.04e-10 800-819
1438	PR01097A	Transient receptor potential family signature I	11.54 5.06e-09 799-820
1439	PR00170A	Voltage-gated Na <sup>+</sup> channel alpha subunit signature I	8.35 3.04e-10 881-900
1439	PR01097A	Transient receptor potential family signature I	11.54 5.06e-09 880-901
1442	IPB000483	Leucine rich repeat C-terminal domain	11.18 6.73e-12 186-200
1442	PR00019A	Leucine-rich repeat signature I	11.72 1.33e-10 106-119
1443	IPB000135A	High mobility group proteins HMG1 and HMG2	11.69 2.69e-31 34-88
1443	IPB000135B	High mobility group proteins HMG1 and HMG2	13.24 3.78e-24 88-132
1443	IPB000203A	GPS domain	18.40 8.43e-09 126-156
1443	IPB000203B	GPS domain	13.98 4.95e-12 286-307
1443	IPB000832C	G-protein coupled receptors family 2 (secretin-like)	19.53 4.64e-14 286-315
1443	PR00249C	Secretin-like GPCR superfamily signature III	15.44 9.56e-10 288-311
1443	PR00249D	Secretin-like GPCR superfamily signature IV	14.10 9.44e-09 327-352
1443	PR00886A	High mobility group (HMG1/HMG2) protein signature I	10.57 8.29e-17 37-59
1443	PR00886B	High mobility group (HMG1/HMG2) protein signature II	10.07 8.97e-13 60-80
1443	PR00886C	High mobility group (HMG1/HMG2) protein signature III	11.20 1.40e-11 104-122
1448	IPB002000D	Lysosome-associated membrane glycoprotein (Lamp)	5.87 5.75e-09 419-432
1448	PR00308C	Type I antifreeze protein signature III	2.79 9.64e-09 758-767
1449	PR00180B	Cellular retinaldehyde-binding protein signature II	15.78 8.03e-09 236-260
1452	IPB001993A	Mitochondrial energy transfer proteins (carrier protein)	16.91 8.65e-10 33-57
1452	IPB001993A	Mitochondrial energy transfer proteins (carrier protein)	16.91 9.39e-09 130-154
1452	IPB001993B	Mitochondrial energy transfer proteins (carrier protein)	9.00 4.21e-09 78-91
1452	IPB001993B	Mitochondrial energy transfer proteins	9.00 7.14e-11 173-186

Table 3B  
424

SEQ ID NO:	Database entry ID	Description	Results*
		(carrier protein)	
1452	PR00926D	Mitochondrial carrier protein signature IV	10.86 7.92e-17 41-59
1452	PR00926E	Mitochondrial carrier protein signature V	11.83 5.74e-19 89-107
1452	PR00926F	Mitochondrial carrier protein signature VI	17.31 1.00e-09 35-57
1452	PR00926F	Mitochondrial carrier protein signature VI	17.31 5.76e-24 132-154
1452	PR00927A	Adenine nucleotide translocator 1 signature I	7.76 8.53e-09 28-40
1452	PR00927B	Adenine nucleotide translocator 1 signature II	15.15 1.44e-12 164-185
1452	PR00927D	Adenine nucleotide translocator 1 signature IV	10.82 3.79e-13 26-39
1452	PR00927E	Adenine nucleotide translocator 1 signature V	14.99 6.32e-11 162-183
1452	PR00927E	Adenine nucleotide translocator 1 signature V	14.99 9.25e-22 67-88
1452	PR00927F	Adenine nucleotide translocator 1 signature VI	12.47 5.50e-14 128-144
1452	PR00927G	Adenine nucleotide translocator 1 signature VII	10.67 5.14e-15 178-193
1454	IPB001993A	Mitochondrial energy transfer proteins (carrier protein)	16.91 8.75e-11 83-107
1454	IPB001993B	Mitochondrial energy transfer proteins (carrier protein)	9.00 1.72e-10 242-255
1455	PR00320A	G protein beta WD-40 repeat signature I	13.15 1.96e-09 128-142
1455	PR00320C	G protein beta WD-40 repeat signature III	12.32 3.67e-09 128-142
1456	IPB000494A	Epidermal growth-factor receptor (EGFR), L domain	22.80 6.59e-25 122-154
1456	IPB000494B	Epidermal growth-factor receptor (EGFR), L domain	24.47 6.00e-18 228-248
1456	IPB000494C	Epidermal growth-factor receptor (EGFR), L domain	24.40 9.45e-33 1099-1145
1456	IPB000494D	Epidermal growth-factor receptor (EGFR), L domain	19.44 1.00e-40 1162-1211
1456	IPB000494E	Epidermal growth-factor receptor (EGFR), L domain	21.65 3.53e-27 1218-1253
1456	IPB001060C	Cell division control protein 15 (CDC15)	26.16 1.00e-09 1183-1202
1456	IPB001245A	Tyrosine kinase catalytic domain	22.45 4.27e-19 1105-1145
1456	IPB001245B	Tyrosine kinase catalytic domain	21.68 6.54e-29 1165-1203
1456	IPB001772C	Kinase associated domain 1	20.66 1.75e-10 1100-1130
1456	IPB002174A	Furin-like cysteine rich region	30.51 4.41e-13 126-157
1456	IPB002174B	Furin-like cysteine rich region	11.91 7.00e-10 1183-1193
1456	IPB003961B	Fibronectin type III domain	12.29 7.43e-11 1184-1194
1456	PR00014D	Fibronectin type III repeat signature IV	15.12 6.09e-09 893-907
1456	PR00109A	Tyrosine kinase catalytic domain signature I	12.56 5.34e-11 1066-1079
1456	PR00109B	Tyrosine kinase catalytic domain signature II	11.07 1.47e-20 1112-1130
1456	PR00109C	Tyrosine kinase catalytic domain signature III	11.86 9.25e-13 1164-1174
1456	PR00109D	Tyrosine kinase catalytic domain signature IV	17.61 4.50e-24 1183-1205
1456	PR00109E	Tyrosine kinase catalytic domain signature V	12.99 4.48e-16 1227-1249
1457	IPB002889B	WSC domain	11.76 8.41e-09 1558-1604
1462	IPB000130	Neutral zinc metalloproteases, zinc-	5.86 4.00e-12 238-248



Table 3B  
425

SEQ ID NO:	Database entry ID	Description	Results*
		binding region	
1462	IPB001818A	Matrixin	14.60 2.13e-29 91-120
1462	IPB001818B	Matrixin	26.48 6.04e-31 137-178
1462	IPB001818C	Matrixin	24.38 4.09e-32 20-65
1462	IPB001818C	Matrixin	24.38 7.43e-35 182-227
1462	IPB001818D	Matrixin	14.91 3.08e-37 232-263
1462	IPB001818E	Matrixin	8.86 7.19e-15 273-286
1462	IPB001818F	Matrixin	11.19 6.36e-13 309-329
1462	PR00138A	Matrixin signature I	12.54 1.64e-16 111-124
1462	PR00138B	Matrixin signature II	14.84 5.21e-10 156-171
1462	PR00138C	Matrixin signature III	20.07 1.78e-16 180-208
1462	PR00138C	Matrixin signature III	20.07 8.38e-15 18-46
1462	PR00138D	Matrixin signature IV	14.57 5.50e-31 238-263
1462	PR00138E	Matrixin signature V	7.10 8.71e-15 273-286
1462	PR00480B	Astacin family signature II	14.35 1.16e-11 233-251
1462	PR00997G	Fragilysin metalloproteinase (M10C) enterotoxin signature VII	11.64 8.01e-10 244-260
1466	IPB000008D	C2 domain	14.83 5.92e-09 288-306
1466	PR01471E	Histamine H3 receptor signature V	5.41 2.43e-10 610-625
1468	IPB000228D	RNA 3'-terminal phosphate cyclase	33.18 4.89e-19 30-67
1468	IPB000228E	RNA 3'-terminal phosphate cyclase	26.21 8.43e-09 105-146
1468	IPB002029A	Aspartate and ornithine carbamoyltransferase family	20.73 7.86e-09 78-105
1470	IPB003874C	CDC45-like protein	5.49 9.16e-10 315-326
1473	IPB002568A	Carlavirus putative nucleic acid binding protein	12.57 9.66e-09 9-34
1473	IPB002999B	Tudor domain	7.50 5.09e-09 29-37
1473	IPB003482B	Transcription factor WhiB	10.09 5.73e-09 10-22
1474	PR01547I	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature IX	15.37 7.94e-28 69-99
1474	PR01547K	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature XI	13.61 7.00e-15 95-114
1474	PR01547L	Saccharomyces cerevisiae 175.8kDa hypothetical protein signature XII	10.35 6.29e-19 119-135
1477	IPB000822	Zinc finger, C2H2 type	14.67 2.88e-09 421-446
1477	PR01407E	Butyrophilin C-terminal DUF signature V	13.29 1.63e-09 56-80
1478	IPB001925C	Eukaryotic porin	23.55 1.84e-28 8-61
1478	PR00185D	Eukaryotic porin signature IV	14.33 9.50e-12 111-128
1482	PR00449E	Transforming protein P21 ras signature V	13.39 1.29e-11 464-486
1489	IPB003859B	Metazoa galactosyltransferase	27.85 8.27e-18 205-257
1489	IPB003859C	Metazoa galactosyltransferase	19.53 3.74e-36 301-335
1489	IPB003859D	Metazoa galactosyltransferase	26.76 6.25e-27 349-382
1491	IPB003191M	Guanylate-binding protein	10.38 7.86e-28 28-58
1491	IPB003191N	Guanylate-binding protein	9.33 6.16e-21 61-91
1491	IPB003191O	Guanylate-binding protein	14.31 7.22e-28 92-132
1491	IPB003191O	Guanylate-binding protein	14.31 8.23e-09 81-121
1493	PR00502B	NUDIX hydrolase family signature II	14.82 4.15e-09 348-363
1494	IPB000117D	Kappa casein	10.18 8.71e-09 117-151
1497	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 3.00e-15 88-125
1498	IPB001395E	Aldo/keto reductase family	20.87 7.39e-23 256-293
1498	IPB001395F	Aldo/keto reductase family	25.03 2.45e-23 317-362
1498	PR00069D	Aldo-keto reductase signature IV	19.00 3.92e-20 260-289
1499	IPB000237B	GRIP domain	30.66 8.24e-09 543-593

Table 3B  
426

SEQ ID NO:	Database entry ID	Description	Results*
1499	IPB000237B	GRIP domain	30.66 8.34e-09 547-597
1499	IPB000996B	Clathrin light chain	20.25 7.55e-09 541-593
1499	IPB001800D	Lipoprotein, type 6	22.73 9.11e-09 558-604
1499	IPB001800D	Lipoprotein, type 6	22.73 9.70e-09 563-609
1499	IPB002101B	MARCKS family	14.58 2.28e-09 545-596
1499	IPB002906A	Ribosomal protein S27a	11.48 3.00e-09 561-593
1499	IPB002906A	Ribosomal protein S27a	11.48 9.81e-10 558-590
1499	IPB003121A	BAF60b domain of the SWIB complex	13.89 9.46e-09 552-577
1499	IPB003134F	Repeat in HS1/Cortactin	15.66 8.49e-09 552-600
1503	PR00049D	Wilm's tumour protein signature IV	0.00 9.57e-10 65-79
1503	PR01217B	Proline rich extensin signature II	4.82 8.55e-09 71-87
1504	IPB000436B	Sushi domain / SCR repeat / CCP module	8.70 8.71e-10 307-318
1504	PR01544D	Arabidopsis thaliana 130.7kDa hypothetical protein signature IV	14.54 9.13e-10 373-394
1511	IPB000822	Zinc finger, C2H2 type	14.67 3.77e-18 343-368
1511	IPB000822	Zinc finger, C2H2 type	14.67 9.25e-19 371-396
1511	IPB001275	DM DNA binding domain	19.17 3.51e-09 331-370
1511	IPB001275	DM DNA binding domain	19.17 5.67e-09 359-398
1511	IPB001909	KRAB box	17.37 4.00e-19 230-264
1511	PR00048A	C2H2-type zinc finger signature I	9.94 5.76e-12 368-381
1511	PR00048A	C2H2-type zinc finger signature I	9.94 7.75e-16 340-353
1511	PR00048B	C2H2-type zinc finger signature II	5.52 7.43e-10 384-393
1517	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-17 1066-1091
1517	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-19 281-306
1517	IPB000822	Zinc finger, C2H2 type	14.67 1.60e-16 393-418
1517	IPB000822	Zinc finger, C2H2 type	14.67 1.60e-16 954-979
1517	IPB000822	Zinc finger, C2H2 type	14.67 1.69e-18 477-502
1517	IPB000822	Zinc finger, C2H2 type	14.67 2.13e-14 1038-1063
1517	IPB000822	Zinc finger, C2H2 type	14.67 4.79e-12 1094-1119
1517	IPB000822	Zinc finger, C2H2 type	14.67 5.50e-19 421-446
1517	IPB000822	Zinc finger, C2H2 type	14.67 5.50e-20 870-895
1517	IPB000822	Zinc finger, C2H2 type	14.67 6.73e-21 926-951
1517	IPB000822	Zinc finger, C2H2 type	14.67 7.00e-20 309-334
1517	IPB000822	Zinc finger, C2H2 type	14.67 8.36e-21 337-362
1517	IPB000822	Zinc finger, C2H2 type	14.67 8.36e-21 982-1007
1517	IPB000822	Zinc finger, C2H2 type	14.67 8.50e-20 449-474
1517	IPB000822	Zinc finger, C2H2 type	14.67 9.25e-20 898-923
1517	IPB000822	Zinc finger, C2H2 type	14.67 9.31e-18 1010-1035
1517	IPB000822	Zinc finger, C2H2 type	14.67 9.40e-16 365-390
1517	IPB000822	Zinc finger, C2H2 type	14.67 9.44e-15 505-530
1517	IPB001275	DM DNA binding domain	19.17 1.12e-10 970-1009
1517	IPB001275	DM DNA binding domain	19.17 3.55e-10 297-336
1517	IPB001275	DM DNA binding domain	19.17 3.69e-12 437-476
1517	IPB001275	DM DNA binding domain	19.17 4.19e-09 886-925
1517	IPB001275	DM DNA binding domain	19.17 4.19e-09 942-981
1517	IPB001275	DM DNA binding domain	19.17 6.71e-13 325-364
1517	IPB001275	DM DNA binding domain	19.17 6.72e-10 1082-1121
1517	IPB001275	DM DNA binding domain	19.17 6.72e-10 914-953
1517	IPB001275	DM DNA binding domain	19.17 8.06e-09 858-897
1517	IPB001275	DM DNA binding domain	19.17 8.18e-10 998-1037
1517	IPB001275	DM DNA binding domain	19.17 9.10e-11 1026-1065
1517	IPB001275	DM DNA binding domain	19.17 9.88e-10 465-504
1517	IPB001909	KRAB box	17.37 8.03e-18 675-709

Table 3B

427

SEQ ID NO:	Database entry ID	Description	Results*
1517	IPB001909	KRAB box	17.37 9.00e-32 76-110
1517	PR00048A	C2H2-type zinc finger signature I	9.94 1.69e-14 306-319
1517	PR00048A	C2H2-type zinc finger signature I	9.94 1.90e-15 446-459
1517	PR00048A	C2H2-type zinc finger signature I	9.94 2.42e-11 474-487
1517	PR00048A	C2H2-type zinc finger signature I	9.94 2.93e-13 502-515
1517	PR00048A	C2H2-type zinc finger signature I	9.94 3.08e-14 1035-1048
1517	PR00048A	C2H2-type zinc finger signature I	9.94 4.38e-09 1007-1020
1517	PR00048A	C2H2-type zinc finger signature I	9.94 5.50e-13 334-347
1517	PR00048A	C2H2-type zinc finger signature I	9.94 5.50e-13 895-908
1517	PR00048A	C2H2-type zinc finger signature I	9.94 6.14e-13 923-936
1517	PR00048A	C2H2-type zinc finger signature I	9.94 7.00e-09 867-880
1517	PR00048A	C2H2-type zinc finger signature I	9.94 7.35e-12 362-375
1517	PR00048A	C2H2-type zinc finger signature I	9.94 7.55e-10 390-403
1517	PR00048A	C2H2-type zinc finger signature I	9.94 8.11e-11 979-992
1517	PR00048A	C2H2-type zinc finger signature I	9.94 8.58e-11 418-431
1517	PR00048A	C2H2-type zinc finger signature I	9.94 9.05e-11 951-964
1517	PR00048A	C2H2-type zinc finger signature I	9.94 9.36e-13 1091-1104
1517	PR00048B	C2H2-type zinc finger signature II	5.52 1.00e-09 911-920
1517	PR00048B	C2H2-type zinc finger signature II	5.52 2.29e-10 939-948
1517	PR00048B	C2H2-type zinc finger signature II	5.52 2.50e-09 995-1004
1517	PR00048B	C2H2-type zinc finger signature II	5.52 2.93e-10 350-359
1517	PR00048B	C2H2-type zinc finger signature II	5.52 3.25e-11 294-303
1517	PR00048B	C2H2-type zinc finger signature II	5.52 3.25e-11 434-443
1517	PR00048B	C2H2-type zinc finger signature II	5.52 3.25e-11 883-892
1517	PR00048B	C2H2-type zinc finger signature II	5.52 3.57e-10 1023-1032
1517	PR00048B	C2H2-type zinc finger signature II	5.52 6.00e-09 462-471
1517	PR00048B	C2H2-type zinc finger signature II	5.52 7.00e-09 967-976
1518	IPB000353B	Class II histocompatibility antigen, beta chain, beta-1 domain	19.16 2.63e-13 21-70
1518	IPB000353C	Class II histocompatibility antigen, beta chain, beta-1 domain	20.11 1.41e-10 71-125
1518	IPB001003B	MHC Class II, alpha chain, alpha-1 domain	14.72 4.45e-09 29-72
1518	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 1.00e-13 42-64
1518	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 9.40e-19 78-115
1518	PR01436D	NADH-ubiquinone oxidoreductase chain 2 signature IV	4.56 2.66e-09 6-18
1521	IPB001618	Coenzyme A transferase	37.23 1.86e-09 56-109
1521	IPB001618	Coenzyme A transferase	37.23 5.75e-32 355-408
1521	IPB001618	Coenzyme A transferase	37.23 6.92e-31 196-249
1521	IPB001839A	Transforming growth factor (TGF) beta family	12.37 3.81e-17 657-673
1521	IPB001839B	Transforming growth factor (TGF) beta family	11.15 2.17e-12 704-718
1521	IPB001839C	Transforming growth factor (TGF) beta family	12.59 2.80e-16 724-741
1521	PR00438A	Growth factor cystine knot domain signature I	11.01 4.00e-09 665-674
1521	PR00669E	Inhibin alpha chain signature V	14.95 5.39e-12 640-657
1523	IPB000034A	Laminin B	22.21 8.56e-09 661-696
1523	IPB001762A	Disintegrin	23.93 4.21e-21 453-493
1523	IPB001762B	Disintegrin	10.06 7.30e-13 500-510
1523	IPB002870A	Reprolysin family propeptide	12.22 6.54e-09 109-125

Table 3B  
428

SEQ ID NO:	Database entry ID	Description	Results*
1523	IPB002870B	Reprolysins family propeptide	24.73 9.38e-26 154-192
1523	IPB002870D	Reprolysins family propeptide	16.31 6.33e-11 323-338
1523	IPB002870E	Reprolysins family propeptide	11.90 7.57e-12 355-367
1523	IPB002870F	Reprolysins family propeptide	18.81 6.85e-19 398-422
1523	IPB003854B	Gibberellin regulated protein	19.26 4.69e-09 470-508
1523	PR00289A	Disintegrin signature I	14.29 2.55e-14 469-488
1523	PR00289B	Disintegrin signature II	11.74 7.27e-10 498-510
1524	IPB000387	Tyrosine specific protein phosphatase and dual specificity protein phosphatase family	10.77 5.00e-11 186-196
1524	IPB001073A	Complement C1q protein	22.14 5.50e-09 249-283
1524	IPB001073A	Complement C1q protein	22.14 5.50e-09 436-470
1524	IPB001695C	Lysyl oxidase	12.69 2.64e-09 391-415
1524	PR00700D	Protein tyrosine phosphatase signature IV	12.83 5.79e-10 183-201
1525	IPB000276C	Rhodopsin-like GPCR superfamily	8.03 4.52e-09 5-16
1525	IPB000276D	Rhodopsin-like GPCR superfamily	9.40 5.24e-14 45-61
1525	PR00237G	Rhodopsin-like GPCR superfamily signature VII	19.23 8.77e-17 35-61
1525	PR00424F	Adenosine receptor signature VI	8.75 5.34e-12 35-45
1525	PR00554F	Adenosine A2B receptor signature VI	8.86 1.75e-16 24-36
1525	PR00554G	Adenosine A2B receptor signature VII	14.18 7.29e-12 69-88
1527	IPB001045B	Spermidine synthase	18.40 1.35e-12 314-356
1528	IPB002048	EF-hand family	7.91 4.32e-09 338-350
1528	PR01286E	Orphan nuclear receptor NOR1 signature V	5.27 6.65e-09 6-27
1528	PR01362H	Flagellar calcium-binding protein (calflagin) signature VIII	4.86 7.13e-09 150-166
1532	IPB001602A	Uncharacterized protein family UPF0047	15.98 8.94e-10 119-137
1534	IPB003861B	E4 protein	9.06 2.96e-09 1051-1065
1534	PR00021A	Small proline-rich protein signature I	3.31 4.55e-09 1044-1056
1535	IPB001359H	Synapsin	22.58 6.59e-09 39-89
1535	PR00671C	Inhibin beta B chain signature III	4.18 9.69e-09 44-63
1538	IPB000095F	PAK-box /P21-Rho-binding	16.47 3.95e-09 146-200
1538	IPB000861G	PKN/rhopilin/rhotekin rho-binding repeat	13.73 9.15e-11 144-193
1538	IPB000959D	POLO box duplicated region	27.01 6.72e-09 196-248
1538	IPB000961D	Protein kinase C-terminal domain	21.23 2.00e-14 138-179
1538	IPB001245A	Tyrosine kinase catalytic domain	22.45 6.63e-16 82-122
1538	IPB001245B	Tyrosine kinase catalytic domain	21.68 7.55e-17 142-180
1538	IPB001772C	Kinase associated domain 1	20.66 9.73e-16 77-107
1538	IPB001772D	Kinase associated domain 1	21.67 4.67e-09 149-188
1538	IPB001772E	Kinase associated domain 1	24.88 2.67e-10 203-242
1538	IPB003527D	MAP kinase	21.53 2.04e-09 135-176
1538	IPB003527G	MAP kinase	17.26 2.38e-09 215-252
1539	PR00502B	NUDIX hydrolase family signature II	14.82 7.53e-09 131-146
1540	IPB000777A	Envelope glycoprotein GP120	12.67 9.67e-09 652-665
1540	IPB000822	Zinc finger, C2H2 type	14.67 9.57e-10 729-754
1540	IPB003036C	Gag P30 core shell protein	11.53 4.86e-09 799-815
1540	PR00049D	Wilm's tumour protein signature IV	0.00 1.79e-10 813-827
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 1.00e-12 537-574
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.52e-11 534-571
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.56e-09 543-580
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 5.85e-11 546-583
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 6.69e-09 555-592
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 7.27e-11 552-589
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 7.36e-10 540-577

Table 3B  
429

SEQ ID NO:	Database entry ID	Description	Results*
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 8.04e-10 528-565
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.12e-10 549-586
1541	IPB000885A	Fibrillar collagen C-terminal domain	11.46 9.41e-10 531-568
1541	IPB000885B	Fibrillar collagen C-terminal domain	19.15 1.77e-10 520-573
1541	IPB000885B	Fibrillar collagen C-terminal domain	19.15 3.70e-09 547-600
1541	IPB000885B	Fibrillar collagen C-terminal domain	19.15 4.11e-09 538-591
1541	IPB000885B	Fibrillar collagen C-terminal domain	19.15 5.01e-09 544-597
1541	IPB000885B	Fibrillar collagen C-terminal domain	19.15 6.49e-10 517-570
1541	IPB000885B	Fibrillar collagen C-terminal domain	19.15 8.11e-10 523-576
1541	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.67e-09 541-594
1541	IPB000885C	Fibrillar collagen C-terminal domain	16.79 1.00e-40 639-678
1541	IPB000885D	Fibrillar collagen C-terminal domain	18.84 4.75e-32 728-765
1541	IPB000885E	Fibrillar collagen C-terminal domain	21.59 1.00e-40 800-841
1541	IPB001073A	Complement C1q protein	22.14 1.00e-08 554-588
1541	IPB001073A	Complement C1q protein	22.14 3.95e-09 551-585
1541	IPB001073A	Complement C1q protein	22.14 7.32e-13 548-582
1541	IPB001073A	Complement C1q protein	22.14 8.03e-09 542-576
1541	IPB001073A	Complement C1q protein	22.14 9.50e-11 545-579
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 2.53e-12 548-600
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.06e-10 551-603
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 5.54e-09 527-579
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.12e-10 545-597
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 8.10e-10 542-594
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 8.59e-09 524-576
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 8.98e-09 536-588
1541	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 9.34e-10 521-573
1541	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 1.12e-09 546-566
1541	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 1.76e-12 549-569
1541	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 1.77e-10 552-572
1542	IPB000033B	Low-density lipoprotein (Ldl) receptor, YWTD repeat	7.05 3.88e-11 1025-1035
1542	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 4.09e-10 927-942
1542	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 8.20e-18 1020-1035
1542	IPB000492B	Protamine 2 (PRM2)	5.26 9.27e-10 430-464
1542	IPB001862F	Membrane attack complex components/perforin/complement C9	29.39 5.39e-13 901-948
1542	IPB001881A	Calcium-binding EGF-like domain	8.72 3.40e-09 916-925
1542	IPB001881B	Calcium-binding EGF-like domain	12.28 7.00e-16 1020-1031
1542	IPB003886D	Extracellular domain in nidogen	13.91 2.77e-10 1020-1039
1542	PR00010A	Type II EGF-like signature I	12.91 2.64e-09 912-923
1542	PR00010C	Type II EGF-like signature III	6.98 1.58e-10 1025-1035

Table 3B  
430

SEQ ID NO:	Database entry ID	Description	Results*
1542	PR00010C	Type II EGF-like signature III	6.98 4.07e-09 932-942
1542	PR00011D	Type III EGF-like signature IV	12.12 2.54e-09 929-947
1542	PR00764F	Complement C9 signature VI	15.74 1.60e-10 918-938
1542	PR00907B	Thrombomodulin signature II	11.50 2.68e-10 1016-1032
1543	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 1.00e-11 88-125
1544	IPB001359H	Synapsin	22.58 2.49e-09 244-294
1544	IPB001394A	Ubiquitin carboxyl-terminal hydrolase family 2	12.64 8.00e-19 478-495
1544	IPB001394B	Ubiquitin carboxyl-terminal hydrolase family 2	17.87 3.16e-19 684-713
1544	IPB001394C	Ubiquitin carboxyl-terminal hydrolase family 2	8.52 1.00e-09 771-780
1544	IPB002000D	Lysosome-associated membrane glycoprotein (Lamp)	5.87 5.58e-09 241-254
1544	IPB002999B	Tudor domain	7.50 7.55e-09 244-252
1544	IPB002999B	Tudor domain	7.50 7.55e-09 245-253
1544	IPB002999B	Tudor domain	7.50 7.55e-09 246-254
1544	IPB003861B	E4 protein	9.06 1.00e-08 246-260
1544	IPB003861B	E4 protein	9.06 9.61e-09 251-265
1544	PR00049D	Wilm's tumour protein signature IV	0.00 1.29e-10 233-247
1544	PR00049D	Wilm's tumour protein signature IV	0.00 2.68e-09 243-257
1544	PR00049D	Wilm's tumour protein signature IV	0.00 4.05e-09 237-251
1544	PR00049D	Wilm's tumour protein signature IV	0.00 4.05e-09 238-252
1544	PR00049D	Wilm's tumour protein signature IV	0.00 4.05e-09 244-258
1544	PR00049D	Wilm's tumour protein signature IV	0.00 5.92e-11 242-256
1544	PR00049D	Wilm's tumour protein signature IV	0.00 6.75e-11 236-250
1544	PR00049D	Wilm's tumour protein signature IV	0.00 7.38e-12 239-253
1544	PR00049D	Wilm's tumour protein signature IV	0.00 8.01e-13 235-249
1544	PR00049D	Wilm's tumour protein signature IV	0.00 8.93e-10 241-255
1544	PR00049D	Wilm's tumour protein signature IV	0.00 9.39e-11 240-254
1544	PR00211B	Glutelin signature II	0.86 7.50e-09 238-258
1544	PR01102M	5-hydroxytryptamine 6 receptor signature XIII	11.13 6.32e-09 229-251
1544	PR01217G	Proline rich extensin signature VII	4.02 5.58e-09 233-258
1544	PR01471E	Histamine H3 receptor signature V	5.41 3.35e-09 235-250
1544	PR01471E	Histamine H3 receptor signature V	5.41 4.23e-09 236-251
1545	IPB001359H	Synapsin	22.58 3.80e-09 277-327
1545	IPB001359H	Synapsin	22.58 8.89e-10 437-487
1545	IPB001359H	Synapsin	22.58 9.83e-09 458-508
1545	IPB002999B	Tudor domain	7.50 2.64e-09 1028-1036
1546	PR00019B	Leucine-rich repeat signature II	11.42 4.33e-10 317-330
1546	PR00416B	Eukaryotic DNA topoisomerase I signature II	15.23 3.03e-09 389-408
1547	IPB000801A	Putative esterase	15.56 7.75e-09 522-533
1547	IPB003533C	Doublecortin	10.25 5.14e-10 627-670
1549	IPB001698A	F-actin capping protein beta subunit	11.94 2.89e-23 47-100
1549	PR00192A	F-actin capping protein beta subunit signature I	8.23 1.47e-27 47-67
1551	IPB002469G	Dipeptidyl peptidase IV, N-terminus	26.76 9.24e-11 362-400
1551	IPB002469H	Dipeptidyl peptidase IV, N-terminus	21.17 6.14e-16 407-442
1551	IPB002469I	Dipeptidyl peptidase IV, N-terminus	10.99 4.86e-16 452-470
1551	IPB002469J	Dipeptidyl peptidase IV, N-terminus	8.97 3.52e-12 534-550
1551	IPB002471B	Prolyl endopeptidase family serine active	24.90 3.66e-11 439-470

Table 3B

431

SEQ ID NO:	Database entry ID	Description	Results*
		site	
1552	IPB000183D	Orn/DAP/Arg decarboxylases family 2	14.78 6.76e-12 38-57
1552	IPB000183E	Orn/DAP/Arg decarboxylases family 2	17.25 2.67e-09 106-129
1552	IPB000183G	Orn/DAP/Arg decarboxylases family 2	21.16 1.00e-17 171-191
1552	IPB003873C	Nonstructural protein NS3/small envelope protein E	7.97 8.03e-10 159-172
1552	PR00163B	Rubredoxin signature II	15.17 4.90e-09 143-159
1552	PR01179E	Ornithine/diaminopimelate/arginine (ODA) decarboxylase family signature V	16.24 5.06e-11 180-193
1552	PR01182G	Ornithine decarboxylase signature VII	11.35 7.00e-14 167-180
1558	IPB001331C	Guanine-nucleotide dissociation stimulators CDC24 family	16.09 2.40e-12 1104-1129
1558	IPB001478B	PDZ domain (also known as DHR or GLGF)	6.12 7.19e-09 198-207
1558	IPB002532J	Hantavirus glycoprotein G2	16.97 8.37e-09 1129-1165
1558	PR00049D	Wilm's tumour protein signature IV	0.00 4.81e-09 1318-1332
1558	PR00049D	Wilm's tumour protein signature IV	0.00 5.73e-09 259-273
1558	PR00554B	Adenosine A2B receptor signature II	12.52 8.85e-09 917-925
1558	PR00834F	HtrA/DegQ protease family signature VI	11.11 5.24e-09 194-206
1559	IPB000095F	PAK-box /P21-Rho-binding	16.47 4.51e-17 374-428
1559	IPB000861D	PKN/rhopilin/rhotekin rho-binding repeat	13.61 6.28e-09 148-184
1559	IPB000961A	Protein kinase C-terminal domain	16.82 4.60e-13 150-184
1559	IPB000961D	Protein kinase C-terminal domain	21.23 5.64e-09 366-407
1559	IPB001245B	Tyrosine kinase catalytic domain	21.68 2.38e-13 370-408
1559	IPB001772A	Kinase associated domain 1	13.64 7.21e-10 148-179
1559	IPB003527A	MAP kinase	17.00 3.49e-09 157-182
1560	IPB003104B	Formin Homology 2 Domain	18.83 6.87e-21 785-814
1560	IPB003104C	Formin Homology 2 Domain	20.33 1.27e-14 957-984
1565	IPB000590A	Hydroxymethylglutaryl-coenzyme A synthase	17.12 1.00e-28 50-74
1565	IPB000590B	Hydroxymethylglutaryl-coenzyme A synthase	16.61 1.67e-37 75-109
1565	IPB000590C	Hydroxymethylglutaryl-coenzyme A synthase	17.39 1.00e-40 110-156
1565	IPB000590D	Hydroxymethylglutaryl-coenzyme A synthase	13.35 5.50e-21 165-196
1565	IPB000590E	Hydroxymethylglutaryl-coenzyme A synthase	20.64 3.08e-31 306-359
1565	IPB000590F	Hydroxymethylglutaryl-coenzyme A synthase	16.80 4.27e-32 384-412
1565	IPB000590G	Hydroxymethylglutaryl-coenzyme A synthase	30.71 1.00e-40 448-502
1568	IPB003016	2-oxo acid dehydrogenases acyltransferase component lipoyl binding site	20.69 2.17e-20 103-137
1572	IPB002043C	Uracil-DNA glycosylase	14.49 2.24e-09 666-676
1574	IPB000637A	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	12.35 7.32e-10 275-295
1574	IPB000840C	Matrix protein (MA)	15.25 8.46e-09 245-290
1574	IPB000949H	ELM2 domain	21.43 9.41e-09 218-264
1574	PR00049D	Wilm's tumour protein signature IV	0.00 4.66e-09 254-268
1574	PR00456E	Ribosomal protein P2 signature V	3.08 5.60e-09 184-198
1575	IPB002567G	Cell fusion glycoprotein K	17.37 9.32e-09 231-270
1577	PR00756B	Membrane alanyl dipeptidase (M1) family signature II	15.53 3.70e-13 232-247

Table 3B

432

SEQ ID NO:	Database entry ID	Description	Results*
1577	PR00756C	Membrane alanyl dipeptidase (M1) family signature III	11.52 1.84e-11 310-320
1579	IPB000221	Protamine P1	5.48 8.01e-11 630-656
1579	IPB000221	Protamine P1	5.48 8.50e-09 626-652
1579	IPB000221	Protamine P1	5.48 9.65e-09 640-666
1579	IPB000492B	Protamine 2 (PRM2)	5.26 5.70e-09 662-696
1579	IPB000492B	Protamine 2 (PRM2)	5.26 7.36e-09 623-657
1579	IPB000822	Zinc finger, C2H2 type	14.67 7.38e-09 586-611
1579	IPB003868A	Herpesvirus UL31-like protein	9.31 6.11e-09 640-654
1580	IPB000215A	Serpins	13.01 5.91e-17 63-86
1580	IPB000215B	Serpins	9.87 2.29e-09 166-178
1580	IPB000215C	Serpins	13.90 4.32e-14 196-210
1580	IPB000215D	Serpins	15.35 4.41e-14 407-433
1581	IPB000557B	Calponin family repeat	17.28 1.63e-11 110-129
1581	IPB001715A	Calponin homology (CH) domain	10.74 7.86e-09 114-124
1581	PR00888C	Smooth muscle protein/calponin family signature III	11.52 9.86e-10 110-125
1581	PR00889B	Calponin signature II	9.54 9.00e-09 121-138
1582	IPB000483	Leucine rich repeat C-terminal domain	11.18 4.54e-10 369-383
1582	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 3.35e-11 475-512
1582	PR00019A	Leucine-rich repeat signature I	11.72 2.09e-09 95-108
1582	PR00019B	Leucine-rich repeat signature II	11.42 3.00e-10 140-153
1582	PR00019B	Leucine-rich repeat signature II	11.42 5.67e-10 92-105
1582	PR00019B	Leucine-rich repeat signature II	11.42 6.45e-09 116-129
1582	PR00019B	Leucine-rich repeat signature II	11.42 9.18e-09 164-177
1586	PR01256B	Otx1 transcription factor signature II	5.92 7.57e-09 33-45
1587	IPB001841	RING finger	10.69 1.00e-09 186-195
1587	IPB002867B	Cysteine-rich domain (C6HC)	22.50 4.15e-15 185-204
1590	IPB002889B	WSC domain	11.76 8.88e-09 98-144
1590	IPB002889B	WSC domain	11.76 9.70e-10 100-146
1592	IPB003888E	FY-rich domain N-terminus	19.80 9.00e-10 661-710
1593	IPB000103A	Pyridine nucleotide-disulphide oxidoreductase class-II	11.55 6.09e-10 747-767
1593	IPB000885A	Fibrillar collagen C-terminal domain	11.46 2.17e-10 481-518
1593	IPB000885A	Fibrillar collagen C-terminal domain	11.46 6.14e-09 475-512
1593	IPB000885B	Fibrillar collagen C-terminal domain	19.15 7.30e-09 609-662
1593	IPB000885B	Fibrillar collagen C-terminal domain	19.15 9.57e-10 320-373
1593	IPB001073A	Complement C1q protein	22.14 4.20e-10 483-517
1593	IPB001073A	Complement C1q protein	22.14 5.50e-09 486-520
1593	IPB001073A	Complement C1q protein	22.14 6.48e-09 342-376
1593	IPB001073A	Complement C1q protein	22.14 7.19e-09 726-760
1593	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 1.00e-08 480-532
1593	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 3.43e-09 468-520
1593	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 4.68e-09 613-665
1593	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 5.21e-10 462-514
1593	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 6.17e-09 483-535
1593	IPB001442A	C-terminal tandem repeated domain in type 4 procollagen	26.12 7.81e-09 459-511



Table 3B

433

SEQ ID NO:	Database entry ID	Description	Results*
1593	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 2.34e-09 487-507
1593	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 6.11e-09 481-501
1593	IPB001442B	C-terminal tandem repeated domain in type 4 procollagen	12.38 7.69e-09 484-504
1593	IPB001451	Bacterial transferase hexapeptide repeat	18.07 2.15e-09 544-579
1593	IPB001451	Bacterial transferase hexapeptide repeat	18.07 5.14e-10 550-585
1593	IPB001451	Bacterial transferase hexapeptide repeat	18.07 8.20e-10 556-591
1593	IPB003474A	GntP family permease	20.01 7.98e-09 179-219
1593	PR00308A	Type I antifreeze protein signature I	3.72 1.00e-08 527-541
1593	PR00308A	Type I antifreeze protein signature I	3.72 3.77e-09 371-385
1593	PR00308A	Type I antifreeze protein signature I	3.72 6.13e-10 372-386
1593	PR00308A	Type I antifreeze protein signature I	3.72 9.31e-09 756-770
1593	PR00308A	Type I antifreeze protein signature I	3.72 9.48e-10 752-766
1593	PR00308C	Type I antifreeze protein signature III	2.79 2.30e-09 375-384
1593	PR00308C	Type I antifreeze protein signature III	2.79 8.82e-09 324-333
1593	PR00308C	Type I antifreeze protein signature III	2.79 9.53e-09 372-381
1593	PR00308C	Type I antifreeze protein signature III	2.79 9.59e-10 373-382
1593	PR00350E	Vitamin D receptor (II nuclear receptor) signature V	11.57 4.21e-09 13-32
1593	PR01286E	Orphan nuclear receptor NOR1 signature V	5.27 9.02e-09 365-386
1593	PR01500B	Tropoelastin signature II	7.08 1.20e-26 229-249
1593	PR01500C	Tropoelastin signature III	6.02 1.31e-24 306-324
1593	PR01511D	Kv1.4 voltage-gated K <sup>+</sup> channel signature IV	3.91 3.47e-09 373-383
1593	PR01511D	Kv1.4 voltage-gated K <sup>+</sup> channel signature IV	3.91 5.85e-09 326-336
1593	PR01511D	Kv1.4 voltage-gated K <sup>+</sup> channel signature IV	3.91 5.85e-09 757-767
1595	IPB000130	Neutral zinc metalloproteinases, zinc-binding region	5.86 5.71e-09 187-197
1595	PR00756C	Membrane alanyl dipeptidase (M1) family signature III	11.52 4.81e-12 151-161
1595	PR00756D	Membrane alanyl dipeptidase (M1) family signature IV	10.78 8.65e-19 187-202
1595	PR00756E	Membrane alanyl dipeptidase (M1) family signature V	10.37 2.42e-15 206-218
1600	IPB000716D	Thyroglobulin type-I repeat	15.49 7.67e-13 96-110
1601	PR00342A	Rhesus blood group protein signature I	13.65 4.75e-21 12-30
1606	IPB002219B	Phorbol esters/diacylglycerol binding domain	12.53 9.13e-09 230-245
1606	IPB002857A	CXXC zinc finger	14.69 4.88e-21 68-99
1606	IPB002999C	Tudor domain	10.33 9.18e-09 540-549
1606	IPB003649B	B-Box C-terminal domain	22.16 7.57e-10 58-77
1606	PR00364D	Disease resistance protein signature IV	14.89 1.00e-08 730-746
1613	IPB002004A	Poly-adenylate binding protein, unique domain	18.93 5.36e-10 150-197
1614	PR00258B	Speract receptor signature II	7.94 7.75e-13 64-75
1614	PR00258C	Speract receptor signature III	9.05 2.72e-09 79-89
1614	PR00258D	Speract receptor signature IV	14.29 6.47e-10 110-124
1614	PR00258E	Speract receptor signature V	14.06 8.78e-13 133-145
1615	IPB003780E	Cytochrome oxidase assembly	15.53 7.40e-14 59-83
1616	IPB000135D	High mobility group proteins HMG1 and	2.13 1.00e-12 318-342

Table 3B

434

SEQ ID NO:	Database entry ID	Description	Results*
		HMG2	
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 1.10e-10 317-341
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 2.16e-15 322-346
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 3.35e-11 327-351
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 3.82e-09 313-337
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 4.18e-09 316-340
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 4.47e-12 319-343
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 4.53e-14 320-344
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 4.58e-11 328-352
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 4.58e-12 326-350
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 5.11e-15 324-348
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 7.00e-09 330-354
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 7.69e-14 325-349
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 7.85e-10 329-353
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 8.12e-16 323-347
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 8.91e-09 315-339
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 9.80e-10 312-336
1616	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 9.87e-15 321-345
1616	IPB000533E	Tropomyosin	11.32 5.16e-09 380-417
1616	IPB000637B	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	14.21 3.45e-09 324-342
1616	IPB000637B	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	14.21 9.18e-09 328-346
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 1.34e-09 322-357
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 2.01e-09 318-353
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 2.80e-09 320-355
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 3.36e-09 321-356
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 4.67e-10 323-358
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 5.86e-10 324-359
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 7.33e-13 319-354
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 7.63e-10 317-352
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 7.64e-09 315-350
1616	IPB001422C	Neuromodulin (GAP-43)	16.82 9.29e-10 314-349
1616	IPB001580F	Calreticulin family	2.93 4.94e-10 325-334
1616	IPB001580F	Calreticulin family	2.93 4.94e-10 326-335
1616	IPB001580F	Calreticulin family	2.93 4.94e-10 327-336
1616	IPB001580F	Calreticulin family	2.93 4.94e-10 328-337

Table 3B  
435

SEQ ID NO:	Database entry ID	Description	Results*
1616	IPB001580F	Calreticulin family	2.93 6.85e-09 331-340
1616	IPB001580F	Calreticulin family	2.93 7.75e-10 336-345
1616	IPB001990C	Granins (chromogranin or secretogranin)	33.59 8.59e-09 309-356
1616	IPB002172	Low density lipoprotein (LDL)-receptor class A (LDLRA) domain	7.37 2.96e-11 52-64
1616	IPB002172	Low density lipoprotein (LDL)-receptor class A (LDLRA) domain	7.37 8.20e-12 93-105
1616	IPB002360C	Involucrin	15.36 9.80e-09 327-368
1616	PR00194D	Tropomyosin signature IV	9.54 8.25e-09 323-346
1616	PR00261A	Low density lipoprotein (LDL) receptor signature I	15.49 4.19e-10 43-64
1616	PR00261A	Low density lipoprotein (LDL) receptor signature I	15.49 4.60e-09 84-105
1616	PR00261B	Low density lipoprotein (LDL) receptor signature II	15.12 2.80e-10 84-105
1616	PR00261B	Low density lipoprotein (LDL) receptor signature II	15.12 6.03e-09 43-64
1616	PR00261C	Low density lipoprotein (LDL) receptor signature III	18.72 3.03e-09 43-64
1616	PR00261D	Low density lipoprotein (LDL) receptor signature IV	16.87 4.25e-10 84-105
1616	PR00261D	Low density lipoprotein (LDL) receptor signature IV	16.87 8.45e-11 43-64
1616	PR00261E	Low density lipoprotein (LDL) receptor signature V	18.62 7.52e-09 84-105
1616	PR00261E	Low density lipoprotein (LDL) receptor signature V	18.62 7.83e-09 43-64
1616	PR00261F	Low density lipoprotein (LDL) receptor signature VI	15.46 2.65e-11 43-64
1616	PR00261F	Low density lipoprotein (LDL) receptor signature VI	15.46 6.91e-10 84-105
1617	IPB002554A	Protein phosphatase 2A regulatory B subunit (B56 family)	19.95 2.29e-27 165-194
1618	IPB003029B	S1 RNA binding domain	9.42 8.80e-09 407-418
1619	IPB000001D	Kringle	11.31 2.88e-12 55-71
1619	IPB000126A	Serine proteases, V8 family	11.75 6.60e-09 55-70
1619	IPB001254A	Serine proteases, trypsin family	9.98 7.65e-14 55-71
1619	PR00722A	Chymotrypsin serine protease family (S1) signature I	12.06 9.36e-14 56-71
1619	PR00839B	V8 serine protease family signature II	11.20 4.95e-09 55-72
1620	IPB000934C	Serine/threonine specific protein phosphatase	20.25 3.00e-34 71-114
1620	IPB000934D	Serine/threonine specific protein phosphatase	14.86 9.55e-17 116-135
1620	IPB000934E	Serine/threonine specific protein phosphatase	25.16 2.88e-27 180-223
1620	PR00114D	Serine/threonine phosphatase family signature IV	11.97 4.60e-24 77-103
1620	PR00114E	Serine/threonine phosphatase family signature V	17.50 7.63e-23 106-133
1620	PR00114F	Serine/threonine phosphatase family signature VI	16.38 2.93e-15 170-190
1620	PR00114G	Serine/threonine phosphatase family signature VII	15.51 9.57e-17 198-214

Table 3B  
436

SEQ ID NO:	Database entry ID	Description	Results*
1623	PR01008D	Flagellar L-ring protein signature IV	9.99 8.13e-17 35-49
1623	PR01008E	Flagellar L-ring protein signature V	13.32 6.48e-17 51-65
1623	PR01010G	Flagellar P-ring protein signature VII	12.40 8.11e-20 88-106
1623	PR01010H	Flagellar P-ring protein signature VIII	12.18 8.88e-18 114-132
1623	PR01010I	Flagellar P-ring protein signature IX	7.74 8.11e-11 178-190
1624	IPB002195C	Dihydroorotase	11.49 8.26e-12 98-109
1624	IPB002604F	Chlorohydrolase	11.98 3.86e-09 130-139
1624	IPB003764D	N-acetylglucosamine-6-phosphate deacetylase	33.54 8.05e-09 114-156
1625	IPB000319C	Aspartate-semialdehyde dehydrogenase	12.28 6.63e-10 61-77
1625	IPB000319D	Aspartate-semialdehyde dehydrogenase	13.92 8.90e-16 91-112
1625	IPB000319E	Aspartate-semialdehyde dehydrogenase	5.68 3.68e-09 155-165
1625	IPB000319F	Aspartate-semialdehyde dehydrogenase	12.86 4.96e-10 238-251
1625	IPB000319G	Aspartate-semialdehyde dehydrogenase	14.74 9.79e-16 258-278
1625	IPB000534A	Semialdehyde dehydrogenase	15.79 7.88e-11 64-81
1628	IPB000847	Bacterial regulatory protein, LysR family	15.44 8.67e-15 21-54
1628	IPB001804D	Isocitrate and isopropylmalate dehydrogenases	24.39 9.18e-09 208-235
1631	IPB000121C	PEP-utilizing enzyme	12.96 5.15e-15 96-108
1631	IPB002192G	Pyruvate phosphate dikinase, PEP/pyruvate binding domain	23.23 3.77e-16 78-109
1631	PR00811C	Bacterial general secretion pathway protein D signature III	10.47 6.20e-09 9-19
1632	IPB002029C	Aspartate and ornithine carbamoyltransferase family	27.21 8.14e-21 401-439
1632	IPB003462C	Ornithine cyclodeaminase/mu-crystallin family	8.98 7.00e-09 500-513
1632	PR00101C	Aspartate carbamoyltransferase signature III	15.49 8.93e-10 410-427
1632	PR01375B	Salmonella/Shigella invasin protein B signature II	7.54 9.67e-09 201-217
1632	PR01507H	Melanin-concentrating hormone 1 receptor signature VIII	11.93 6.53e-09 354-373
1634	IPB001568B	Ribonuclease T2 family	11.65 1.00e-17 101-112
1643	IPB001014	Ribosomal L23 protein	22.21 6.03e-20 217-249
1645	PR01303D	Plasmodium circumsporozoite protein signature IV	10.57 6.88e-09 486-503
1647	IPB001518B	Argininosuccinate synthase	12.16 6.87e-12 58-70
1647	IPB001518C	Argininosuccinate synthase	27.71 3.57e-27 72-114
1647	IPB001518E	Argininosuccinate synthase	20.16 5.24e-14 172-188
1651	PR00453A	Von Willebrand factor type A domain signature I	11.78 3.89e-13 88-105
1651	PR00453B	Von Willebrand factor type A domain signature II	13.84 6.10e-13 125-139
1651	PR00453C	Von Willebrand factor type A domain signature III	11.84 3.57e-10 246-254
1654	IPB000222B	Protein phosphatase 2C subfamily	15.80 6.66e-09 94-104
1654	IPB000222C	Protein phosphatase 2C subfamily	6.84 5.86e-11 118-127
1657	IPB000917A	Sulfatase	9.52 4.79e-10 47-58
1658	IPB000063	Thioredoxin	9.89 1.69e-16 346-359
1658	IPB000063	Thioredoxin	9.89 2.38e-16 481-494
1658	PR00421A	Thioredoxin family signature I	10.35 1.00e-10 481-489
1658	PR00421A	Thioredoxin family signature I	10.35 2.13e-11 346-354
1658	PR00421B	Thioredoxin family signature II	10.23 1.00e-11 354-363
1658	PR00421B	Thioredoxin family signature II	10.23 7.75e-12 489-498

Table 3B

437

SEQ ID NO:	Database entry ID	Description	Results*
1658	PR00421C	Thioredoxin family signature III	11.02 1.00e-09 394-405
1658	PR00421C	Thioredoxin family signature III	11.02 3.84e-10 533-544
1661	IPB003134F	Repeat in HS1/Cortactin	15.66 9.68e-09 101-149
1662	IPB002360C	Involucrin	15.36 9.00e-09 56-97
1662	PR01270A	Histone deacetylase superfamily signature I	14.72 6.85e-10 1235-1258
1662	PR01270B	Histone deacetylase superfamily signature II	9.70 5.76e-18 1269-1284
1662	PR01270C	Histone deacetylase superfamily signature III	10.89 3.57e-09 1360-1370
1663	PR01104A	Anaphylatoxin chemotactic receptor signature I	13.83 8.50e-09 83-94
1666	IPB000737	Serine protease inhibitor, squash family	13.14 8.31e-14 85-104
1666	IPB002098	Seminal vesicle protein I repeats	9.00 1.00e-40 31-67
1666	IPB002098	Seminal vesicle protein I repeats	9.00 1.43e-16 43-79
1666	IPB002098	Seminal vesicle protein I repeats	9.00 2.54e-17 19-55
1666	IPB002098	Seminal vesicle protein I repeats	9.00 3.79e-15 37-73
1666	IPB002098	Seminal vesicle protein I repeats	9.00 7.15e-18 25-61
1666	IPB002098	Seminal vesicle protein I repeats	9.00 9.83e-10 49-85
1666	IPB002221B	WAP-type (Whey Acidic Protein) four-disulfide core domain	17.12 3.14e-20 92-113
1666	PR00003A	4-disulphide core signature I	14.31 7.00e-09 69-78
1666	PR00003C	4-disulphide core signature III	7.81 7.63e-12 98-107
1666	PR00003D	4-disulphide core signature IV	9.63 8.50e-09 108-116
1667	IPB001909	KRAB box	17.37 1.00e-32 108-142
1673	IPB003150C	RFX DNA-binding domain	25.34 7.08e-09 248-294
1674	IPB001322A	Intermediate filament tail domain	30.52 7.88e-11 63-116
1674	PR01002A	Flagellar protein FlgJ signature I	10.02 9.65e-09 261-281
1676	IPB001865A	Ribosomal protein S2	35.08 8.94e-11 53-106
1678	PR00795C	Ryanodine receptor signature III	6.10 7.92e-09 107-131
1684	IPB001998A	Xylose isomerase	15.52 3.05e-09 214-227
1684	IPB002198	Short-chain dehydrogenase/reductase (SDR) superfamily	19.72 8.80e-20 133-168
1684	PR00080C	Short-chain dehydrogenase/reductase (SDR) superfamily signature III	14.49 8.83e-13 153-172
1684	PR00081A	Glucose/ribitol dehydrogenase family signature I	10.07 8.20e-10 5-22
1684	PR00081C	Glucose/ribitol dehydrogenase family signature III	12.71 7.97e-09 127-143
1684	PR00081D	Glucose/ribitol dehydrogenase family signature IV	15.22 4.71e-10 153-172
1684	PR01167F	Insect alcohol dehydrogenase family signature VI	11.22 2.34e-09 152-170
1685	IPB001526C	Ly-6/u-PAR domain	13.04 2.03e-09 197-212
1685	PR01514F	Kv2.1 voltage-gated K <sup>+</sup> channel signature VI	2.00 5.29e-09 220-231
1686	IPB000095E	PAK-box /P21-Rho-binding	17.62 7.35e-13 679-724
1686	IPB000961C	Protein kinase C-terminal domain	15.48 5.50e-12 678-712
1686	IPB001245A	Tyrosine kinase catalytic domain	22.45 8.88e-16 671-711
1686	IPB001245B	Tyrosine kinase catalytic domain	21.68 8.64e-09 736-774
1686	IPB003527C	MAP kinase	14.70 5.30e-09 663-711
1686	IPB003605C	GS motif preceding kinase domain in TGF beta receptor	14.92 2.91e-15 570-617
1686	IPB003605D	GS motif preceding kinase domain in TGF beta receptor	12.41 4.39e-16 667-708

Table 3B

438

SEQ ID NO:	Database entry ID	Description	Results*
1686	IPB003605E	GS motif preceding kinase domain in TGF beta receptor	21.14 3.61e-15 710-749
1686	IPB003605F	GS motif preceding kinase domain in TGF beta receptor	23.21 2.56e-12 826-879
1686	PR00653D	Activin type II receptor signature IV	12.01 9.76e-12 697-718
1686	PR00653E	Activin type II receptor signature V	14.33 2.93e-16 871-890
1687	IPB000832A	G-protein coupled receptors family 2 (secretin-like)	13.81 3.25e-13 94-109
1687	PR00491A	Vasoactive intestinal peptide receptor signature I	12.62 2.98e-09 109-120
1687	PR01154A	Vasoactive intestinal peptide receptor 1 signature I	11.05 3.84e-29 58-79
1687	PR01154B	Vasoactive intestinal peptide receptor 1 signature II	12.24 1.39e-22 95-111
1688	PR00401A	SH2 domain signature I	12.88 6.29e-15 214-228
1688	PR00452A	SH3 domain signature I	9.57 2.35e-10 150-160
1688	PR00452B	SH3 domain signature II	11.47 4.91e-12 164-179
1688	PR00452C	SH3 domain signature III	10.63 4.60e-09 181-190
1688	PR00499D	Neutrophil cytosol factor 2 signature IV	11.47 8.29e-10 152-172
1688	PR01260A	Na <sup>+</sup> /Ca <sup>2+</sup> exchanger isoform 1 signature I	9.46 2.75e-09 5-17
1693	PR00019A	Leucine-rich repeat signature I	11.72 1.27e-09 660-673
1693	PR00019A	Leucine-rich repeat signature I	11.72 1.55e-09 614-627
1693	PR00019B	Leucine-rich repeat signature II	11.42 7.55e-11 527-540
1693	PR00019B	Leucine-rich repeat signature II	11.42 8.91e-09 611-624
1694	PR00499B	Neutrophil cytosol factor 2 signature II	9.48 4.73e-09 34-53
1695	IPB001359H	Synapsin	22.58 4.58e-09 150-200
1704	IPB000272	ATP1G1/PLM/MAT8 family	14.36 8.53e-09 28-63
1705	IPB000892B	Ribosomal protein S26E	13.49 3.50e-27 160-196
1706	IPB000906D	ZU5 domain	23.89 2.54e-09 368-422
1706	IPB000906D	ZU5 domain	23.89 5.71e-09 437-491
1706	IPB000906D	ZU5 domain	23.89 8.54e-09 401-455
1706	IPB000906E	ZU5 domain	22.11 1.38e-09 448-488
1706	IPB000906G	ZU5 domain	25.85 3.64e-10 382-430
1706	IPB000906G	ZU5 domain	25.85 5.87e-09 451-499
1706	IPB001580F	Calreticulin family	2.93 8.20e-09 317-326
1706	IPB003716C	RNA polymerase omega subunit	13.72 9.39e-09 314-334
1706	PR00019B	Leucine-rich repeat signature II	11.42 1.55e-09 742-755
1706	PR00019B	Leucine-rich repeat signature II	11.42 5.67e-10 773-786
1706	PR01415A	Ankyrin repeat signature I	12.73 1.45e-13 415-427
1706	PR01415A	Ankyrin repeat signature I	12.73 4.19e-10 451-463
1706	PR01415B	Ankyrin repeat signature II	10.23 1.90e-10 463-475
1706	PR01415B	Ankyrin repeat signature II	10.23 2.93e-12 427-439
1708	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-18 235-260
1708	IPB000822	Zinc finger, C2H2 type	14.67 2.93e-17 207-232
1708	IPB000822	Zinc finger, C2H2 type	14.67 6.57e-10 263-288
1708	IPB000822	Zinc finger, C2H2 type	14.67 8.71e-10 53-78
1708	IPB001275	DM DNA binding domain	19.17 7.04e-09 223-262
1708	PR00048A	C2H2-type zinc finger signature I	9.94 1.00e-08 204-217
1708	PR00048A	C2H2-type zinc finger signature I	9.94 4.38e-09 232-245
1708	PR00048A	C2H2-type zinc finger signature I	9.94 7.00e-09 260-273
1708	PR00048A	C2H2-type zinc finger signature I	9.94 9.63e-09 50-63
1708	PR00048B	C2H2-type zinc finger signature II	5.52 7.43e-10 220-229
1708	PR00498A	Neutrophil cytosol factor 1 signature I	12.92 6.74e-09 353-361
1709	IPB002870A	Reprolysins family propeptide	12.22 4.09e-10 108-124

Table 3B  
439

SEQ ID NO:	Database entry ID	Description	Results*
1710	IPB000222B	Protein phosphatase 2C subfamily	15.80 9.49e-09 1539-1549
1710	IPB001389D	Flocculin repeat	9.66 5.93e-09 1712-1739
1710	PR01391C	Binary toxin B family signature III	14.87 5.44e-09 259-277
1711	PR01099B	Hydroxyethylthiazole kinase family signature II	14.42 5.44e-09 6-27
1712	IPB000716C	Thyroglobulin type-1 repeat	17.62 5.41e-12 287-305
1712	IPB000867A	Insulin-like growth factor-binding protein	7.19 7.19e-09 48-55
1712	IPB000867B	Insulin-like growth factor-binding protein	11.44 4.13e-18 66-82
1712	PR01169I	Ceratitis capitata alcohol dehydrogenase signature IX	10.85 9.41e-09 425-445
1715	IPB003308C	Integrase zinc-binding domain	19.91 3.42e-33 22-75
1715	IPB003308D	Integrase zinc-binding domain	8.41 7.43e-09 94-106
1715	IPB003308E	Integrase zinc-binding domain	9.08 1.20e-11 166-179
1717	IPB001134C	Netrin, C-terminus	17.82 7.82e-10 1317-1331
1717	IPB001599A	Alpha-2-macroglobulin family	10.97 1.45e-21 175-193
1717	IPB001599B	Alpha-2-macroglobulin family	7.45 3.35e-14 264-276
1717	IPB001599C	Alpha-2-macroglobulin family	14.40 4.60e-15 291-307
1717	IPB001599D	Alpha-2-macroglobulin family	11.61 4.94e-15 783-793
1717	IPB001599E	Alpha-2-macroglobulin family	11.06 1.60e-10 810-819
1717	IPB001599F	Alpha-2-macroglobulin family	18.95 5.30e-27 840-869
1717	IPB001599G	Alpha-2-macroglobulin family	13.87 1.53e-13 988-997
1717	IPB001599H	Alpha-2-macroglobulin family	18.42 1.00e-23 1020-1047
1717	IPB001599I	Alpha-2-macroglobulin family	10.83 4.33e-12 1057-1066
1717	IPB001599J	Alpha-2-macroglobulin family	20.99 3.10e-16 1088-1113
1717	IPB001599K	Alpha-2-macroglobulin family	8.15 7.60e-13 1273-1284
1717	IPB001599L	Alpha-2-macroglobulin family	18.66 4.00e-25 1304-1331
1717	IPB001599M	Alpha-2-macroglobulin family	13.29 1.75e-12 1445-1456
1717	IPB001599N	Alpha-2-macroglobulin family	24.85 9.22e-26 1498-1530
1717	IPB001983D	Translationally controlled tumor protein	22.98 6.63e-13 1213-1253
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 1.00e-12 217-241
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 2.41e-15 219-243
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 2.86e-17 220-244
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 3.43e-14 218-242
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 4.58e-11 214-238
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 5.23e-12 223-247
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 5.60e-10 213-237
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 6.18e-09 212-236
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 7.56e-15 222-246
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 8.06e-11 224-248
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 8.30e-14 216-240
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 9.13e-12 215-239
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 9.14e-17 221-245

Table 3B

440

SEQ ID NO:	Database entry ID	Description	Results*
		HMG2	
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 9.18e-11 225-249
1718	IPB000135D	High mobility group proteins HMG1 and HMG2	2.13 9.64e-09 226-250
1718	IPB000637B	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	14.21 1.00e-08 224-242
1718	IPB000637B	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	14.21 1.00e-08 225-243
1718	IPB000637B	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	14.21 2.09e-09 223-241
1718	IPB000637B	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	14.21 4.82e-09 221-239
1718	IPB000637B	HMG-I and HMG-Y DNA-binding domain (A+T-hook)	14.21 8.36e-09 222-240
1718	IPB001422C	Neuromodulin (GAP-43)	16.82 2.07e-10 218-253
1718	IPB001422C	Neuromodulin (GAP-43)	16.82 2.13e-09 222-257
1718	IPB001422C	Neuromodulin (GAP-43)	16.82 3.70e-09 220-255
1718	IPB001422C	Neuromodulin (GAP-43)	16.82 3.84e-10 217-252
1718	IPB001422C	Neuromodulin (GAP-43)	16.82 6.68e-10 223-258
1718	IPB001422C	Neuromodulin (GAP-43)	16.82 6.74e-09 221-256
1718	IPB001422C	Neuromodulin (GAP-43)	16.82 8.93e-10 225-260
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 224-233
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 225-234
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 226-235
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 227-236
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 228-237
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 229-238
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 230-239
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 231-240
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 232-241
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 233-242
1718	IPB001580F	Calreticulin family	2.93 4.94e-10 234-243
1718	IPB001580F	Calreticulin family	2.93 7.75e-09 235-244
1718	IPB002126A	Cadherin domain	14.68 7.00e-09 119-135
1718	IPB002126B	Cadherin domain	12.04 5.20e-16 422-439
1718	IPB002126B	Cadherin domain	12.04 7.86e-13 156-173
1718	IPB002126C	Cadherin domain	12.35 8.41e-09 398-408
1718	PR00194D	Tropomyosin signature IV	9.54 1.94e-09 222-245
1718	PR00194D	Tropomyosin signature IV	9.54 2.05e-09 220-243
1718	PR00205A	Cadherin signature I	17.38 3.50e-13 108-127
1718	PR00205B	Cadherin signature II	20.09 3.30e-11 551-580
1718	PR00205B	Cadherin signature II	20.09 4.66e-14 434-463
1718	PR00205B	Cadherin signature II	20.09 9.50e-21 168-197
1718	PR00205C	Cadherin signature III	13.59 3.25e-11 397-409
1718	PR00205D	Cadherin signature IV	12.22 4.46e-19 413-432
1718	PR00205E	Cadherin signature V	10.82 4.32e-11 432-445
1718	PR00205F	Cadherin signature VI	19.57 4.77e-10 380-406
1718	PR00205F	Cadherin signature VI	19.57 5.20e-13 493-519
1718	PR00205F	Cadherin signature VI	19.57 7.30e-13 115-141
1718	PR00205G	Cadherin signature VII	13.05 8.71e-14 416-433
1720	IPB003360A	US22-like viral protein	16.29 4.27e-09 133-146
1722	IPB001128	none	11.42 5.15e-13 469-481
1722	PR00385A	P450 superfamily signature I	13.94 4.19e-10 341-358



Table 3B

441

SEQ ID NO:	Database entry ID	Description	Results*
1722	PR00385C	P450 superfamily signature III	17.08 6.29e-10 395-406
1722	PR00385D	P450 superfamily signature IV	11.11 9.47e-10 470-479
1722	PR00385E	P450 superfamily signature V	13.50 5.50e-09 479-490
1722	PR00463H	E-class P450 group I signature VIII	11.44 5.68e-10 469-479
1722	PR00465B	E-class P450 group IV signature II	17.22 7.92e-10 113-136
1722	PR00465C	E-class P450 group IV signature III	20.35 8.92e-23 332-358
1722	PR00465D	E-class P450 group IV signature IV	15.21 8.62e-16 390-406
1722	PR00465E	E-class P450 group IV signature V	15.83 4.41e-13 423-437
1722	PR00465F	E-class P450 group IV signature VI	13.57 1.00e-13 439-457
1722	PR00465G	E-class P450 group IV signature VII	15.73 8.07e-16 463-479
1722	PR00465H	E-class P450 group IV signature VIII	19.55 9.25e-21 479-497
1726	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 7.65e-12 91-113
1726	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 2.80e-10 130-167
1727	IPB001285A	Synaptophysin/synaptoporin	12.19 1.00e-40 26-72
1727	IPB001285B	Synaptophysin/synaptoporin	21.42 2.00e-35 73-104
1727	IPB001285B	Synaptophysin/synaptoporin	21.42 7.86e-12 72-103
1727	IPB001285C	Synaptophysin/synaptoporin	15.24 1.00e-40 108-148
1727	IPB001285D	Synaptophysin/synaptoporin	12.74 1.00e-40 149-183
1727	IPB001285E	Synaptophysin/synaptoporin	9.55 1.00e-40 184-230
1727	IPB001285F	Synaptophysin/synaptoporin	6.39 1.00e-40 231-275
1727	IPB001285F	Synaptophysin/synaptoporin	6.39 1.54e-09 248-292
1727	IPB001285F	Synaptophysin/synaptoporin	6.39 3.46e-09 242-286
1727	IPB001285F	Synaptophysin/synaptoporin	6.39 8.65e-11 258-302
1727	IPB001285F	Synaptophysin/synaptoporin	6.39 9.83e-11 263-307
1727	PR00220A	Synaptophysin/synaptoporin family signature I	13.69 1.41e-27 23-45
1727	PR00220B	Synaptophysin/synaptoporin family signature II	17.03 8.31e-28 47-72
1727	PR00220C	Synaptophysin/synaptoporin family signature III	11.86 7.00e-28 100-124
1727	PR00220D	Synaptophysin/synaptoporin family signature IV	10.60 8.13e-28 132-155
1727	PR00220E	Synaptophysin/synaptoporin family signature V	4.56 5.68e-23 199-217
1727	PR00761A	Bindin precursor signature I	6.20 7.00e-10 287-303
1728	IPB002360C	Involucrin	15.36 1.40e-09 114-155
1728	IPB002360C	Involucrin	15.36 7.80e-09 105-146
1728	IPB002558F	I/LWEQ domain	14.64 7.33e-09 155-185
1729	IPB003137	Protease associated (PA) domain	22.40 2.86e-15 551-581
1729	PR00747C	Glycosyl hydrolase family 47 signature III	12.34 5.14e-11 145-163
1729	PR00747E	Glycosyl hydrolase family 47 signature V	14.81 2.08e-18 225-242
1729	PR00747H	Glycosyl hydrolase family 47 signature VIII	12.76 5.50e-17 326-346
1730	PR00209C	Alpha/beta gliadin family signature III	3.37 5.39e-09 77-90
1731	IPB000276B	Rhodopsin-like GPCR superfamily	4.97 9.05e-09 192-203
1731	IPB000276C	Rhodopsin-like GPCR superfamily	8.03 3.50e-11 244-255
1731	IPB000276D	Rhodopsin-like GPCR superfamily	9.40 9.47e-14 285-301
1731	PR00237A	Rhodopsin-like GPCR superfamily signature I	9.81 2.42e-13 35-59
1731	PR00237B	Rhodopsin-like GPCR superfamily signature II	12.45 5.24e-16 68-89
1731	PR00237D	Rhodopsin-like GPCR superfamily signature IV	9.76 1.47e-11 133-154

Table 3B

442

SEQ ID NO:	Database entry ID	Description	Results*
1731	PR00237E	Rhodopsin-like GPCR superfamily signature V	13.03 8.07e-11 184-207
1731	PR00237F	Rhodopsin-like GPCR superfamily signature VI	14.34 2.89e-17 234-258
1731	PR00237G	Rhodopsin-like GPCR superfamily signature VII	19.23 8.11e-19 275-301
1731	PR00244A	Neurokinin receptor signature I	10.06 5.00e-15 29-40
1731	PR00244B	Neurokinin receptor signature II	12.18 1.00e-13 55-64
1731	PR00244C	Neurokinin receptor signature III	13.48 2.87e-12 108-118
1731	PR00244D	Neurokinin receptor signature IV	7.47 9.36e-14 125-135
1731	PR00244E	Neurokinin receptor signature V	5.67 1.44e-11 183-192
1731	PR00244F	Neurokinin receptor signature VI	12.03 6.29e-16 204-217
1731	PR00244G	Neurokinin receptor signature VII	9.65 1.45e-15 268-279
1731	PR00244H	Neurokinin receptor signature VIII	13.18 5.50e-26 294-312
1731	PR01025A	Neurokinin NK2 receptor signature I	8.63 5.50e-16 2-15
1731	PR01025B	Neurokinin NK2 receptor signature II	12.20 7.11e-17 15-28
1731	PR01025C	Neurokinin NK2 receptor signature III	18.53 2.23e-16 154-166
1731	PR01025D	Neurokinin NK2 receptor signature IV	12.91 5.60e-18 168-182
1731	PR01025E	Neurokinin NK2 receptor signature V	8.25 7.23e-25 212-230
1731	PR01025F	Neurokinin NK2 receptor signature VI	14.52 1.00e-17 257-269
1731	PR01025G	Neurokinin NK2 receptor signature VII	9.85 6.65e-23 312-331
1733	IPB000917A	Sulfatase	9.52 5.26e-10 44-55
1734	PR00464A	Group II E-class P450 signature I	19.27 7.26e-14 167-187
1735	IPB000237B	GRIP domain	30.66 8.14e-09 273-323
1735	IPB002097C	Profilin	22.96 4.04e-10 986-1040
1735	IPB002097C	Profilin	22.96 5.08e-09 991-1045
1735	IPB003753G	Exonuclease VII, large subunit	31.72 8.85e-09 787-831
1735	PR00392C	Profilin signature III	14.42 2.29e-09 986-1006
1735	PR00392D	Profilin signature IV	10.64 4.35e-09 1008-1022
1738	IPB001138	Fungal transcriptional regulatory protein, N-terminus	12.17 9.79e-09 62-78
1739	IPB001509A	NAD dependent epimerase/dehydratase family	14.43 9.65e-09 379-400
1741	IPB003644J	Domain found in Na-Ca exchanger and integrin-beta4	13.64 5.67e-09 57-111
1741	PR00915G	Luteovirus group 1 coat protein signature VII	15.74 6.58e-10 303-324
1745	IPB001283A	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7	16.26 6.67e-10 29-48
1745	IPB001283B	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7	13.09 2.33e-12 57-72
1745	IPB001283E	Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7	13.34 5.50e-20 127-147
1745	IPB003106E	Homeobox associated leucine zipper	13.73 6.90e-09 25-55
1745	PR00837A	Allergen V5/Tpx-1 family signature I	14.69 1.56e-14 57-75
1745	PR00837C	Allergen V5/Tpx-1 family signature III	16.31 5.50e-20 126-142
1745	PR00838D	Venom allergen 5 signature IV	8.59 3.57e-09 57-75
1745	PR00838G	Venom allergen 5 signature VII	15.39 8.60e-18 125-144
1749	PR00672D	Inhibin beta C chain signature IV	10.52 6.40e-09 96-112
1749	PR01035I	Tetracycline resistance protein signature IX	13.24 3.86e-09 568-589
1749	PR01221G	Rodent urinary protein signature VII	12.11 8.68e-10 61-78
1750	IPB000885A	Fibrillar collagen C-terminal domain	11.46 8.16e-09 1852-1889
1751	IPB001991A	Sodium:dicarboxylate symporter family	10.78 2.53e-20 9-40
1751	IPB001991B	Sodium:dicarboxylate symporter family	14.40 2.29e-33 324-355

Table 3B

443

SEQ ID NO:	Database entry ID	Description	Results*
1751	IPB001991C	Sodium:dicarboxylate symporter family	9.91 3.40e-24 375-394
1751	PR00173B	Glutamate-aspartate symporter signature II	7.82 7.87e-12 19-39
1751	PR00173C	Glutamate-aspartate symporter signature III	10.67 9.44e-26 327-352
1753	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 6.54e-09 286-323
1755	IPB002617C	Vesicular monoamine transporter	16.23 5.03e-10 91-143
1755	PR01035E	Tetracycline resistance protein signature V	10.87 2.08e-09 101-123
1755	PR01035F	Tetracycline resistance protein signature VI	12.88 7.52e-09 332-352
1757	IPB000300B	Inositol polyphosphate related phosphatase family	15.64 1.93e-16 468-491
1757	IPB000300C	Inositol polyphosphate related phosphatase family	14.50 8.50e-12 538-548
1757	IPB000300D	Inositol polyphosphate related phosphatase family	29.87 4.51e-18 589-623
1757	IPB000300E	Inositol polyphosphate related phosphatase family	15.50 1.43e-14 774-790
1757	IPB000492B	Protamine 2 (PRM2)	5.26 4.23e-09 83-117
1758	PR00169B	Potassium channel signature II	16.49 8.80e-28 239-267
1758	PR00169C	Potassium channel signature III	17.53 9.83e-11 307-330
1758	PR01496C	Shaker voltage-gated K <sup>+</sup> channel family signature III	9.37 6.92e-12 202-213
1758	PR01496D	Shaker voltage-gated K <sup>+</sup> channel family signature IV	9.74 3.25e-21 236-252
1758	PR01496E	Shaker voltage-gated K <sup>+</sup> channel family signature V	10.86 2.11e-14 259-270
1758	PR01509A	Kv1.2 voltage-gated K <sup>+</sup> channel signature I	6.59 3.57e-17 278-292
1758	PR01509B	Kv1.2 voltage-gated K <sup>+</sup> channel signature II	8.19 1.26e-17 294-306
1759	IPB000631B	Uncharacterized protein family UPF0031	10.96 5.00e-12 100-112
1759	IPB000631C	Uncharacterized protein family UPF0031	12.87 7.75e-12 376-387
1759	IPB000631E	Uncharacterized protein family UPF0031	7.82 9.25e-11 458-468
1762	IPB001734A	Sodium:solute symporter family	14.69 8.83e-13 40-56
1762	IPB001734B	Sodium:solute symporter family	18.26 5.50e-26 155-195
1762	IPB001734C	Sodium:solute symporter family	18.21 2.16e-21 448-477
1763	IPB000483	Leucine rich repeat C-terminal domain	11.18 2.23e-12 358-372
1763	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 5.15e-09 462-499
1763	PR00019A	Leucine-rich repeat signature I	11.72 5.50e-12 86-99
1763	PR00019A	Leucine-rich repeat signature I	11.72 8.91e-09 278-291
1763	PR00019B	Leucine-rich repeat signature II	11.42 1.00e-08 107-120
1763	PR00019B	Leucine-rich repeat signature II	11.42 8.77e-11 83-96
1763	PR00019B	Leucine-rich repeat signature II	11.42 9.67e-10 275-288
1764	IPB000130	Neutral zinc metallopeptidases, zinc-binding region	5.86 6.57e-09 294-304
1764	IPB001762A	Disintegrin	23.93 5.21e-10 241-281
1764	IPB001862A	Membrane attack complex components/perforin/complement C9	12.54 1.76e-10 416-431
1764	IPB002861B	Reeler domain	10.50 1.41e-09 422-450
1764	IPB002870D	Reprolysin family propeptide	16.31 6.25e-12 265-280
1764	IPB002870E	Reprolysin family propeptide	11.90 4.41e-12 296-308
1764	IPB002870F	Reprolysin family propeptide	18.81 3.73e-13 335-359
1764	PR00977A	Scytalidopepsin B aspartic protease (A4) signature I	7.30 9.11e-09 233-255
1765	PR00824A	Hepatic lipase signature I	8.11 4.44e-22 6-24

Table 3B  
444

SEQ ID NO:	Database entry ID	Description	Results*
1768	IPB000337A	G-protein coupled receptors family 3 (Metabotropic glutamate receptor-like)	9.17 4.32e-14 47-58
1768	IPB000337B	G-protein coupled receptors family 3 (Metabotropic glutamate receptor-like)	24.51 9.90e-09 150-198
1768	IPB001828A	Receptor family ligand binding region	13.21 9.05e-09 74-86
1768	PR00248A	Metabotropic glutamate GPCR signature I	7.46 9.10e-16 46-58
1768	PR00248B	Metabotropic glutamate GPCR signature II	12.69 2.89e-17 75-90
1768	PR00248C	Metabotropic glutamate GPCR signature III	11.96 8.20e-24 90-109
1768	PR00248D	Metabotropic glutamate GPCR signature IV	14.23 9.00e-15 153-179
1768	PR01054A	Metabotropic glutamate receptor 4 signature I	7.72 1.22e-25 9-26
1768	PR01054B	Metabotropic glutamate receptor 4 signature II	9.34 1.47e-26 26-44
1768	PR01054C	Metabotropic glutamate receptor 4 signature III	10.37 1.15e-22 131-147
1769	IPB000008B	C2 domain	17.91 3.02e-10 226-243
1769	IPB000008C	C2 domain	23.37 2.80e-31 400-439
1769	IPB000008C	C2 domain	23.37 3.04e-10 265-304
1769	IPB000008D	C2 domain	14.83 8.59e-14 455-473
1769	PR00360A	C2 domain signature I	15.18 3.00e-10 398-410
1769	PR00360B	C2 domain signature II	11.64 2.33e-12 427-440
1769	PR00360B	C2 domain signature II	11.64 8.88e-13 292-305
1769	PR00399A	Synaptotagmin signature I	15.05 5.50e-11 236-251
1769	PR00399B	Synaptotagmin signature II	14.30 5.50e-10 251-264
1769	PR00399B	Synaptotagmin signature II	14.30 7.92e-12 385-398
1769	PR00399C	Synaptotagmin signature III	15.89 1.63e-10 308-323
1769	PR00399D	Synaptotagmin signature IV	12.72 6.33e-10 328-338
1770	IPB000534A	Semialdehyde dehydrogenase	15.79 4.21e-09 287-304
1771	IPB001195	Glycophorin A	7.52 6.07e-10 717-745
1771	IPB001359H	Synapsin	22.58 7.99e-09 871-921
1771	IPB001514D	RNA polymerases D/30 to 40 Kd subunits	19.99 9.43e-09 1326-1352
1771	IPB002126A	Cadherin domain	14.68 4.15e-11 412-428
1771	IPB002126B	Cadherin domain	12.04 1.50e-09 232-249
1771	IPB002126B	Cadherin domain	12.04 3.25e-12 448-465
1771	IPB002126B	Cadherin domain	12.04 7.25e-09 123-140
1771	PR00205A	Cadherin signature I	17.38 5.66e-10 293-312
1771	PR00205A	Cadherin signature I	17.38 7.84e-11 75-94
1771	PR00205B	Cadherin signature II	20.09 2.73e-09 135-164
1771	PR00205B	Cadherin signature II	20.09 4.06e-10 460-489
1771	PR00205B	Cadherin signature II	20.09 8.80e-15 244-273
1771	PR00205C	Cadherin signature III	13.59 7.75e-09 317-329
1771	PR00205D	Cadherin signature IV	12.22 3.42e-14 439-458
1771	PR00205D	Cadherin signature IV	12.22 3.83e-12 223-242
1771	PR00205D	Cadherin signature IV	12.22 6.40e-12 544-563
1771	PR00205D	Cadherin signature IV	12.22 6.71e-11 331-350
1771	PR00205D	Cadherin signature IV	12.22 7.51e-10 114-133
1771	PR00205E	Cadherin signature V	10.82 4.86e-09 458-471
1771	PR00205F	Cadherin signature VI	19.57 1.95e-11 408-434
1771	PR00205F	Cadherin signature VI	19.57 2.50e-15 513-539
1771	PR00205F	Cadherin signature VI	19.57 3.65e-12 300-326
1771	PR00205G	Cadherin signature VII	13.05 2.65e-09 334-351
1771	PR00205G	Cadherin signature VII	13.05 4.65e-11 226-243
1771	PR00205G	Cadherin signature VII	13.05 7.91e-10 442-459
1771	PR00205G	Cadherin signature VII	13.05 9.57e-14 547-564

Table 3B  
445

SEQ ID NO:	Database entry ID	Description	Results*
1771	PR01157D	P2 purinoceptor signature IV	16.03 5.42e-09 815-827
1771	PR01383B	Claudin-10 signature II	5.36 5.05e-09 724-736
1774	IPB000301B	Transmembrane 4 family	18.74 4.68e-26 492-527
1774	IPB000301D	Transmembrane 4 family	26.02 7.30e-23 662-702
1774	PR00259B	Transmembrane four family signature II	13.09 3.03e-17 486-512
1774	PR00259C	Transmembrane four family signature III	16.39 5.76e-19 513-541
1774	PR00259D	Transmembrane four family signature IV	14.90 8.54e-18 676-702
1775	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 1.95e-10 111-122
1775	IPB000276A	Rhodopsin-like GPCR superfamily	11.56 5.50e-09 199-210
1775	PR00237E	Rhodopsin-like GPCR superfamily signature V	13.03 3.90e-10 280-303
1775	PR00245A	Olfactory receptor signature I	10.98 1.90e-10 85-96
1775	PR00245A	Olfactory receptor signature I	10.98 2.50e-10 173-184
1775	PR00245B	Olfactory receptor signature II	13.73 1.00e-10 210-222
1775	PR00245B	Olfactory receptor signature II	13.73 8.71e-09 122-134
1775	PR00245C	Olfactory receptor signature III	14.65 4.91e-15 257-273
1775	PR00245D	Olfactory receptor signature IV	9.34 8.41e-13 317-326
1775	PR00534A	Melanocortin receptor family signature I	12.77 9.56e-11 44-56
1778	IPB000920C	Myelin P0 protein	15.78 2.33e-09 160-212
1778	IPB001388	Synaptobrevin	26.97 7.60e-30 115-169
1778	PR00219A	Synaptobrevin signature I	10.78 1.35e-11 118-137
1778	PR00219B	Synaptobrevin signature II	9.91 5.30e-09 138-157
1778	PR00219C	Synaptobrevin signature III	6.91 2.20e-13 20-39
1780	IPB000353B	Class II histocompatibility antigen, beta chain, beta-1 domain	19.16 4.96e-12 177-226
1780	IPB000353C	Class II histocompatibility antigen, beta chain, beta-1 domain	20.11 9.18e-10 228-282
1780	IPB001039A	Major histocompatibility complex protein, Class I	17.17 1.00e-40 15-68
1780	IPB001039C	Major histocompatibility complex protein, Class I	19.82 8.08e-12 151-204
1780	IPB001039D	Major histocompatibility complex protein, Class I	16.49 1.00e-40 229-283
1780	IPB003006A	Immunoglobulin and major histocompatibility complex domain	17.51 3.25e-18 198-220
1780	IPB003006B	Immunoglobulin and major histocompatibility complex domain	20.23 5.50e-28 235-272
1780	IPB003363E	Glycoprotein GG/GX	13.35 2.18e-10 282-314
1782	IPB000084A	PE N-terminus	14.90 3.44e-09 644-691
1782	IPB000258G	Bacterial ice-nucleation proteins octamer repeat	8.61 4.96e-09 681-734
1782	IPB000258H	Bacterial ice-nucleation proteins octamer repeat	10.13 6.03e-09 719-773
1782	IPB000765	GTP1/OBG family	26.91 9.84e-10 449-492
1782	IPB000790B	ATP synthase alpha subunit, C-terminal	10.45 9.64e-10 424-467
1782	IPB001140A	ABC transporter transmembrane region	21.73 2.42e-18 437-483
1782	IPB001140B	ABC transporter transmembrane region	15.62 7.70e-12 555-593
1782	IPB001482B	Bacterial type II secretion system protein E	12.05 3.70e-09 447-469
1782	PR00300A	ATP-dependent Clp protease ATP-binding subunit signature I	7.82 8.27e-09 451-469
1782	PR00449A	Transforming protein P21 ras signature I	12.48 1.67e-09 449-470
1782	PR00671B	Inhibin beta B chain signature II	4.29 9.41e-09 396-415
1783	IPB000175A	Sodium:neurotransmitter symporter family	16.29 1.00e-40 211-260
1783	IPB000175B	Sodium:neurotransmitter symporter family	19.12 3.37e-32 298-332

Table 3B

446

SEQ ID NO:	Database entry ID	Description	Results*
1783	IPB000175C	Sodium:neurotransmitter symporter family	15.09 3.75e-13 385-436
1783	IPB000175D	Sodium:neurotransmitter symporter family	23.45 1.00e-40 451-503
1783	IPB000175E	Sodium:neurotransmitter symporter family	21.88 1.67e-30 544-583
1783	IPB000175F	Sodium:neurotransmitter symporter family	25.63 1.00e-30 639-678
1783	IPB000175G	Sodium:neurotransmitter symporter family	16.18 6.32e-20 700-722
1783	PR00176A	Sodium/chloride neurotransmitter symporter signature I	16.97 5.50e-24 211-232
1783	PR00176B	Sodium/chloride neurotransmitter symporter signature II	7.07 2.13e-22 240-259
1783	PR00176C	Sodium/chloride neurotransmitter symporter signature III	10.57 3.16e-23 283-309
1783	PR00176D	Sodium/chloride neurotransmitter symporter signature IV	8.96 7.92e-21 412-429
1783	PR00176E	Sodium/chloride neurotransmitter symporter signature V	11.14 3.84e-19 494-514
1783	PR00176F	Sodium/chloride neurotransmitter symporter signature VI	11.11 9.57e-20 548-567
1783	PR00176G	Sodium/chloride neurotransmitter symporter signature VII	13.12 7.67e-16 630-650
1783	PR00176H	Sodium/chloride neurotransmitter symporter signature VIII	15.94 4.52e-16 670-690
1783	PR01204A	Glycine neurotransmitter transporter type 1 (GLYT-1) signature I	15.11 3.25e-18 187-200
1783	PR01204B	Glycine neurotransmitter transporter type 1 (GLYT-1) signature II	12.89 1.37e-15 201-211
1783	PR01204C	Glycine neurotransmitter transporter type 1 (GLYT-1) signature III	12.34 1.60e-16 335-347
1783	PR01204D	Glycine neurotransmitter transporter type 1 (GLYT-1) signature IV	10.51 6.86e-18 349-364
1784	IPB001627K	Sema domain	13.76 3.02e-09 662-674
1784	IPB002165D	Plexin repeat	14.72 1.15e-10 662-674
1784	PR01319A	Glial cell line-derived neurotrophic factor receptor alpha 3 signature I	3.85 7.08e-09 27-39
1785	IPB000647A	CTF/NF-I family	11.22 1.00e-40 644-693
1785	IPB000647B	CTF/NF-I family	10.06 2.14e-30 694-718
1785	IPB000647C	CTF/NF-I family	9.27 1.00e-40 724-765
1785	IPB000647D	CTF/NF-I family	12.61 1.20e-32 766-790
1785	IPB000647E	CTF/NF-I family	12.21 9.18e-38 791-821
1786	IPB000483	Leucine rich repeat C-terminal domain	11.18 4.54e-10 675-689
1786	PR00019A	Leucine-rich repeat signature I	11.72 4.67e-10 243-256
1786	PR00019A	Leucine-rich repeat signature I	11.72 6.45e-09 434-447
1786	PR00019B	Leucine-rich repeat signature II	11.42 4.27e-11 431-444
1786	PR00019B	Leucine-rich repeat signature II	11.42 5.09e-09 479-492
1786	PR00019B	Leucine-rich repeat signature II	11.42 7.55e-09 597-610
1786	PR00019B	Leucine-rich repeat signature II	11.42 8.64e-09 240-253
1786	PR00364D	Disease resistance protein signature IV	14.89 4.86e-09 188-204
1789	PR00702F	Acriflavin resistance protein family signature VI	14.03 1.46e-17 49-72
1789	PR00702G	Acriflavin resistance protein family signature VII	10.67 6.00e-21 74-97
1789	PR00702H	Acriflavin resistance protein family signature VIII	14.64 3.67e-14 153-170
1789	PR00702I	Acriflavin resistance protein family signature IX	16.75 1.39e-15 216-230

Table 3B

447

SEQ ID NO:	Database entry ID	Description	Results*
1790	PR00797F	Streptopain (C10) cysteine protease family signature VI	11.89 8.81e-09 41-62
1791	IPB000849D	GlpT family of transporters	22.39 9.25e-31 55-91
1791	IPB000849E	GlpT family of transporters	19.05 2.88e-36 108-145
1791	IPB000849F	GlpT family of transporters	15.06 8.00e-20 162-179
1792	IPB000917C	Sulfatase	13.32 4.67e-11 593-603
1792	IPB003661A	His Kinase A domain	9.35 1.60e-09 244-253
1796	IPB000203A	GPS domain	18.40 6.63e-12 903-933
1796	IPB000203B	GPS domain	13.98 9.12e-12 988-1009
1796	IPB000483	Leucine rich repeat C-terminal domain	11.18 6.04e-11 371-385
1796	IPB000832C	G-protein coupled receptors family 2 (secretin-like)	19.53 7.33e-09 988-1017
1796	PR00019A	Leucine-rich repeat signature I	11.72 1.00e-09 315-328
1796	PR00019B	Leucine-rich repeat signature II	11.42 8.00e-10 312-325
1798	IPB002208C	SecY protein	15.51 4.15e-17 30-50
1798	IPB002208D	SecY protein	22.30 2.29e-36 82-118
1798	PR00303C	Preprotein translocase SecY subunit signature III	15.05 7.84e-19 34-57
1798	PR00303D	Preprotein translocase SecY subunit signature IV	15.48 1.95e-22 72-97
1798	PR00303E	Preprotein translocase SecY subunit signature V	14.43 7.19e-23 98-121
1798	PR00825B	Vespid venom allergen phospholipase A1 signature II	13.74 9.40e-09 234-254
1800	PR00081A	Glucose/ribitol dehydrogenase family signature I	10.07 9.70e-12 179-196
1800	PR00669B	Inhibin alpha chain signature II	5.88 3.93e-09 80-96
1801	IPB000109C	PTR peptide transporters (PTR2)	8.21 3.74e-13 12-24
1801	IPB000109D	PTR peptide transporters (PTR2)	25.09 1.00e-29 464-511
1802	IPB000033B	Low-density lipoprotein (ldl) receptor, YWTD repeat	7.05 4.96e-11 494-504
1802	IPB000152	Aspartic acid and asparagine hydroxylation site	8.86 1.47e-14 489-504
1802	IPB001169K	Integrin beta, C-terminus	27.45 2.90e-10 970-1012
1802	IPB001774C	Delta serrate ligand	18.25 2.19e-09 973-1015
1802	IPB001881B	Calcium-binding EGF-like domain	12.28 6.68e-11 489-500
1802	IPB001969A	Eukaryotic and viral aspartic protease active site	16.37 7.00e-10 378-394
1802	PR00010C	Type II EGF-like signature III	6.98 7.10e-10 494-504
1802	PR00792B	Pepsin (A1) aspartic protease family signature II	12.65 3.25e-13 331-344
1802	PR00792C	Pepsin (A1) aspartic protease family signature III	8.65 9.31e-13 380-391
1803	IPB000729A	PMP-22/EMP/MP20 family	19.70 1.95e-09 1-22
1803	IPB000729A	PMP-22/EMP/MP20 family	19.70 2.20e-10 18-39
1803	IPB000729B	PMP-22/EMP/MP20 family	13.56 5.68e-10 49-59
1803	IPB000729C	PMP-22/EMP/MP20 family	37.83 3.47e-23 87-139
1803	IPB000729C	PMP-22/EMP/MP20 family	37.83 7.83e-22 80-132
1803	IPB000729D	PMP-22/EMP/MP20 family	18.96 4.94e-18 160-187
1803	PR01077B	Claudin signature II	14.12 8.80e-10 49-55
1803	PR01077C	Claudin signature III	13.60 9.43e-12 63-73
1804	IPB001717I	Anion exchanger family	7.53 2.07e-11 18-58
1804	IPB001717L	Anion exchanger family	14.31 9.28e-14 84-130
1804	IPB001717M	Anion exchanger family	19.47 6.23e-32 133-185

Table 3B

448

SEQ ID NO:	Database entry ID	Description	Results*
1804	IPB001717N	Anion exchanger family	14.40 1.00e-30 183-208
1804	IPB001717O	Anion exchanger family	10.14 4.76e-16 209-236
1804	IPB001717Q	Anion exchanger family	17.70 5.77e-09 232-273
1804	PR00165F	Anion exchanger family signature VI	14.73 1.50e-11 42-49
1804	PR00165J	Anion exchanger family signature X	6.53 4.18e-09 127-135
1804	PR00165K	Anion exchanger family signature XI	12.74 6.40e-11 156-164
1804	PR00165M	Anion exchanger family signature XIII	9.29 7.26e-12 255-265
1804	PR01120D	Plant CLC chloride channel signature IV	10.53 6.46e-09 134-141
1804	PR01231E	HCO <sub>3</sub> - transporter superfamily signature V	8.87 4.38e-15 40-52
1804	PR01231H	HCO <sub>3</sub> - transporter superfamily signature VIII	12.16 9.18e-15 111-124
1805	IPB001394A	Ubiquitin carboxyl-terminal hydrolase family 2	12.64 3.16e-16 786-803
1805	IPB001394B	Ubiquitin carboxyl-terminal hydrolase family 2	17.87 3.05e-21 1215-1244
1805	IPB001394C	Ubiquitin carboxyl-terminal hydrolase family 2	8.52 3.50e-10 1290-1299
1805	IPB001394D	Ubiquitin carboxyl-terminal hydrolase family 2	9.19 4.60e-10 1308-1317
1806	PR01130C	Delayed-early response protein/equilibrative nucleoside transporter signature III	7.16 6.00e-09 187-210
1806	PR01130E	Delayed-early response protein/equilibrative nucleoside transporter signature V	14.23 6.07e-10 309-330
1806	PR01130F	Delayed-early response protein/equilibrative nucleoside transporter signature VI	5.34 3.54e-11 337-354
1807	IPB001919B	Cellulose-binding domain, bacterial type	14.22 2.97e-09 104-128
1810	IPB003662A	General substrate transporters	18.97 8.67e-14 736-768
1811	PR01228C	Eggshell protein signature III	5.69 3.90e-09 51-66
1813	IPB000260A	NADH-ubiquinone oxidoreductase chain 4, amino terminus	12.52 2.80e-09 16-49
1813	IPB000260A	NADH-ubiquinone oxidoreductase chain 4, amino terminus	12.52 4.66e-10 19-52
1813	IPB000444B	Xanthine/uracil permeases family	22.44 4.21e-09 51-94
1813	IPB000444B	Xanthine/uracil permeases family	22.44 4.32e-09 43-86
1813	IPB000684L	Eukaryotic RNA polymerase II heptapeptide repeat	3.49 7.25e-09 841-883
1813	IPB000822	Zinc finger, C2H2 type	14.67 1.00e-18 272-297
1813	IPB000822	Zinc finger, C2H2 type	14.67 2.71e-10 416-441
1813	IPB000822	Zinc finger, C2H2 type	14.67 7.19e-14 487-512
1813	IPB000822	Zinc finger, C2H2 type	14.67 8.00e-13 515-540
1813	IPB000822	Zinc finger, C2H2 type	14.67 8.11e-12 300-325
1813	IPB000822	Zinc finger, C2H2 type	14.67 9.05e-12 773-798
1813	IPB001457	NADH-ubiquinone/plastoquinone oxidoreductase chain 6	19.26 5.81e-09 5-46
1813	PR00048A	C2H2-type zinc finger signature I	9.94 6.25e-09 770-783
1813	PR00048A	C2H2-type zinc finger signature I	9.94 9.53e-11 269-282
1813	PR00048B	C2H2-type zinc finger signature II	5.52 6.79e-10 285-294
1813	PR00219C	Synaptobrevin signature III	6.91 6.34e-09 33-52
1815	IPB000001D	Kringle	11.31 2.02e-13 53-69
1815	IPB000001H	Kringle	12.24 5.29e-11 318-328
1815	IPB000126A	Serine proteases, V8 family	11.75 5.07e-11 53-68



Table 3B  
449

SEQ ID NO:	Database entry ID	Description	Results*
1815	IPB000177M	Apple domain	9.18 8.55e-16 226-260
1815	IPB000177N	Apple domain	10.17 5.21e-15 308-342
1815	IPB001254A	Serine proteases, trypsin family	9.98 1.00e-15 53-69
1815	IPB001254B	Serine proteases, trypsin family	15.01 6.29e-18 316-339
1815	PR00722A	Chymotrypsin serine protease family (S1) signature I	12.06 9.28e-15 54-69
1815	PR00722C	Chymotrypsin serine protease family (S1) signature III	10.74 8.88e-13 315-327
1815	PR00839B	V8 serine protease family signature II	11.20 6.42e-09 53-70
1817	IPB003819A	Taurine catabolism dioxygenase TauD/TfdA	10.95 3.03e-09 200-207
1817	IPB003819B	Taurine catabolism dioxygenase TauD/TfdA	21.71 5.64e-26 222-255
1817	IPB003819C	Taurine catabolism dioxygenase TauD/TfdA	11.87 2.74e-17 288-304
1817	IPB003819D	Taurine catabolism dioxygenase TauD/TfdA	27.59 4.68e-38 319-359
1819	IPB000387	Tyrosine specific protein phosphatase and dual specificity protein phosphatase family	10.77 2.38e-13 1434-1444
1819	PR00014C	Fibronectin type III repeat signature III	14.47 4.00e-09 599-617
1819	PR00014C	Fibronectin type III repeat signature III	14.47 4.86e-09 688-706
1819	PR00014C	Fibronectin type III repeat signature III	14.47 5.00e-10 1056-1074
1819	PR00014C	Fibronectin type III repeat signature III	14.47 7.00e-09 866-884
1819	PR00213E	Myelin P0 protein signature V	5.51 1.66e-09 1186-1210
1819	PR00700A	Protein tyrosine phosphatase signature I	6.05 5.80e-10 1306-1313
1819	PR00700C	Protein tyrosine phosphatase signature III	13.89 7.88e-17 1390-1407
1819	PR00700D	Protein tyrosine phosphatase signature IV	12.83 3.86e-20 1431-1449
1819	PR00700F	Protein tyrosine phosphatase signature VI	10.33 4.71e-12 1478-1488
1819	PR01371E	Salmonella/Yersinia modular tyrosine phosphatase signature V	11.71 5.98e-12 1432-1443
1820	IPB003842A	Vacuolating cytotoxin	7.27 9.28e-09 100-146
1821	IPB000276D	Rhodopsin-like GPCR superfamily	9.40 9.65e-09 405-421
1821	PR00245D	Olfactory receptor signature IV	9.34 1.53e-13 359-368
1821	PR00245E	Olfactory receptor signature V	8.96 6.88e-13 406-417
1822	IPB003861B	E4 protein	9.06 1.00e-08 34-48
1822	PR00494E	Fanconi anaemia group C protein signature V	10.42 8.26e-09 621-640
1822	PR01039A	Tryptophanyl-tRNA synthetase signature I	14.70 9.59e-09 200-216

\* Results include in order: accession number subtype; raw score; p-value; position of signature in amino acid sequence

Table 4A

450

SEQ ID	Model	Description	E-value	Score	Repeats	Position
912	MIF	Macrophage migration inhibitory factor (MIF)	2.4e-54	194.0	1	124-227
912	GST_C	Glutathione S-transferase, C-terminal domain	6.2e-25	92.6	2	1-45:410-525
912	GST_N	Glutathione S-transferase, N-terminal domain	1.4e-16	62.5	1	324-400
913	tsp_1	Thrombospondin type 1 domain	2.9e-23	90.7	2	124-174:181-231
913	EGF	EGF-like domain	1.6e-20	81.6	7	474-508:514-553:559-591:597-633:692-720:802-836:842-882
914	MIF	Macrophage migration inhibitory factor (MIF)	4.5e-50	179.8	1	2-101
915	MIF	Macrophage migration inhibitory factor (MIF)	6.1e-68	239.1	1	2-129
916	trypsin	Trypsin	3.8e-99	314.8	1	445-675
916	kringle	Kringle domain	4.2e-45	163.3	1	286-367
916	fn2	Fibronectin type II domain	4.5e-35	94.9	1	108-148
916	EGF	EGF-like domain	6.2e-15	63.0	2	164-197:245-278
916	fn1	Fibronectin type I domain	5.2e-11	45.0	1	202-237
917	laminin_EGF	Laminin EGF-like (Domains III and V)	2.3e-158	539.5	12	282-346:349-409:412-469:472-520:523-565:786-831:834-877:880-925:928-984:987-1036:1039-1094:1095-1139
917	laminin_Nterm	Laminin N-terminal (Domain VI)	5.9e-105	362.1	1	45-280
918	EGF	EGF-like domain	3.2e-53	190.2	9	33-68:74-110:116-151:161-197:208-243:247-282:288-323:329-362:368-404
918	CUB	CUB domain	1.1e-16	68.8	1	708-817
919	Parathyroid	Parathyroid hormone family	3.1e-82	286.6	1	46-175
920	EGF	EGF-like domain	2.2e-22	87.8	9	2-36:47-82:88-122:128-163:232-263:276-306:319-349:353-394:407-439
921	IL8	Small cytokines (intecrine/chemokine), inter	2.7e-09	35.4	1	25-65
922	Zn_carbOpept	Zinc carboxypeptidase	3.1e-119	409.5	1	149-429
922	Propep_M14	Carboxypeptidase activation peptide	5.6e-20	79.8	1	7-84
923	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	1.8e-67	237.6	1	62-248
927	FKBP	FKBP-type peptidyl-prolyl cis-trans isomeras	1.3e-44	149.6	2	40-111:112-122
930	FAD_binding_4	FAD binding domain	6.5e-89	308.8	2	11-176:186-381
930	FAD-oxidase_C	FAD linked oxidases, C-terminal domain	2.6e-63	223.8	1	383-647
933	Fringe	Fringe-like	1.4e-06	-12.8	1	262-487

Table 4A

451

SEQ ID	Model	Description	E-value	Score	Repeats	Position
937	14-3-3	14-3-3 protein	1.1e-124	427.6	1	58-282
938	S_100	S-100/ICaBP type calcium binding domain	3.6e-23	90.4	1	65-108
938	efhand	EF hand	0.0012	25.5	1	114-142
940	ig	Immunoglobulin domain	1.2e-33	113.5	5	38-105:139-206:241-293:326-377:412-487
941	LRR	Leucine Rich Repeat	3.6e-30	113.7	8	42-63:64-87:88-112:113-136:137-160:161-184:185-208:209-232
941	LRRCT	Leucine rich repeat C-terminal domain	1.2e-07	38.9	1	242-292
943	AMP-binding	AMP-binding enzyme	1.2e-16	62.6	1	82-159
944	serpin	Serpin (serine protease inhibitor)	7e-134	455.1	1	173-532
947	RUN	RUN domain	8e-44	159.0	1	31-163
947	FYVE	FYVE zinc finger	1.6e-18	70.7	1	529-576
950	lectin_c	Lectin C-type domain	2.4e-08	41.1	1	130-234
952	Glyco_transf_8	Glycosyl transferase family 8	0.01	-47.7	1	1255-1526
953	TPR	TPR Domain	0.0044	23.7	1	171-204
956	Branch	Core-2/I-Branching enzyme	5.5e-102	352.2	1	74-306
957	fn3	Fibronectin type III domain	1.1e-15	65.6	1	30-108
958	Peptidase_M10	Matrixin	5.9e-121	415.2	1	58-225
958	hemopexin	Hemopexin	2e-58	207.5	4	305-347:349-393:398-445:447-487
962	chromo	'chromo' (CHRmatin Organization MODifier)	5.5e-09	35.7	1	64-84
964	ig	Immunoglobulin domain	1.8e-05	22.6	1	34-115
967	7tm_1	7 transmembrane receptor (rhodopsin family)	2.1e-26	86.0	2	41-154:161-237
968	adh_zinc	Zinc-binding dehydrogenases	7.4e-60	212.3	1	179-495
970	E1_dehydrog	Dehydrogenase E1 component	2.1e-183	622.7	1	129-425
971	ig	Immunoglobulin domain	9.3e-11	39.7	1	45-122
973	zf-MIZ	MIZ zinc finger	3.2e-32	120.5	1	279-331
973	SAP	SAP domain	7.4e-06	32.9	1	2-36
975	WD40	WD domain, G-beta repeat	5e-08	40.1	3	234-270:447-483:491-528
977	An_peroxidase	Animal haem peroxidase	5.6e-140	478.4	1	137-701
978	Defensin_propep	Defensin propeptide	1.9e-17	71.4	1	32-82
978	defensins	Mammalian defensin	8.6e-05	23.3	1	95-123
981	ig	Immunoglobulin domain	3.6e-06	24.9	1	49-130
982	ig	Immunoglobulin domain	1.2e-09	36.1	1	36-112
985	tsp_1	Thrombospondin type 1 domain	0.0075	17.1	1	151-206
987	LRR	Leucine Rich Repeat	1.1e-58	208.4	12	78-101:102-125:126-149:150-173:174-197:198-221:222-245:246-269:270-293:294-317:318-341:342-

Table 4A

452

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						365
988	fer4_NifH	4Fe-4S iron sulfur cluster binding proteins	0.0015	15.9	1	110-134
988	ParA	ParA family ATPase	0.006	-0.3	1	177-283
989	vwd	von Willebrand factor type D domain	1.6e-37	134.9	1	364-514
989	vwc	von Willebrand factor type C domain	5.3e-35	129.7	5	50-105:108-163:166-224:238-289:301-357
989	TIL	Trypsin Inhibitor like cysteine rich domain	8.9e-06	32.6	1	629-682
992	ig	Immunoglobulin domain	9.6e-17	59.0	2	34-111:150-220
994	ig	Immunoglobulin domain	1.6e-28	97.0	3	43-110:143-197:230-291
995	SCP	SCP-like extracellular protein	4.7e-41	149.8	1	4-173
996	ig	Immunoglobulin domain	8.9e-06	23.7	2	73-123:153-206
999	RNase_PH	3' exoribonuclease family	1.5e-28	108.3	1	264-469
1000	ig	Immunoglobulin domain	2.6e-05	22.2	1	38-102
1002	zf-C2H2	Zinc finger, C2H2 type	7.6e-54	192.3	7	183-205:211-233:239-261:267-289:295-317:323-345:351-373
1003	zf-C2H2	Zinc finger, C2H2 type	2.1e-114	393.5	16	260-282:297-319:325-347:353-375:381-403:409-431:437-459:465-487:493-515:521-543:549-571:577-599:605-627:633-655:661-683:689-711
1007	LIM	LIM domain	1.5e-37	138.2	3	390-448:450-507:510-576
1010	zf-C2H2	Zinc finger, C2H2 type	4.2e-06	33.7	3	2-24:29-52:154-177
1011	60s_ribosomal	60s Acidic ribosomal protein	1.1e-28	94.4	2	1-22:25-89
1012	abhydrolase_2	Phospholipase/Carboxylesterase	1.3e-24	95.2	1	9-213
1016	Anti_proliferat	BTG1 family	3.6e-73	256.5	1	11-253
1017	zf-C2H2	Zinc finger, C2H2 type	9.3e-71	248.5	11	147-169:213-231:269-291:325-347:353-375:381-403:409-431:437-459:465-487:493-515:521-543
1017	KRAB	KRAB box	1.7e-21	84.8	1	42-122
1020	LIM	LIM domain	7e-36	132.6	4	78-136:139-197:200-256:259-318
1024	zf-C2H2	Zinc finger, C2H2 type	3.3e-53	190.2	10	156-183:184-206:212-234:240-262:268-290:296-318:324-346:352-374:380-402:408-

Table 4A

453

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						430
1024	KRAB	KRAB box	2.1e-42	154.3	1	15-77
1025	RAG2	Recombination activating protein 2	0	1380.2	1	1-527
1026	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	3.7e-09	33.9	1	16-54
1029	Keratin_B2	Keratin, high sulfur B2 protein	3.8e-06	16.4	2	5-149:150-281
1031	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-tr	1.5e-106	367.4	1	5-165
1035	zf-C2H2	Zinc finger, C2H2 type	9.9e-114	391.2	17	100-122:132-154:160-182:188-210:216-238:244-266:272-294:300-322:328-350:356-378:384-406:412-434:440-462:468-490:496-518:524-546:552-574
1035	KRAB	KRAB box	2.3e-16	67.8	1	4-54
1038	Sulfate_transp	Sulfate transporter family	4.3e-103	355.9	2	2-284:441-751
1038	STAS	STAS domain	4.8e-20	80.0	1	774-987
1039	ubiquitin	Ubiquitin family	1.4e-09	39.6	1	55-126
1042	ig	Immunoglobulin domain	1.2e-30	103.8	5	62-129:163-229:264-316:349-400:433-501
1043	UPAR_LY6	u-PAR/Ly-6 domain	9.1e-59	208.7	1	63-190
1047	MHC_I	Class I Histocompatibility antigen, domains	3.2e-147	502.5	1	25-203
1047	ig	Immunoglobulin domain	0.057	11.4	1	220-285
1050	ig	Immunoglobulin domain	2.3e-09	35.2	1	38-110
1051	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.0019	-4.7	1	4-194
1054	SRCR	Scavenger receptor cysteine-rich domain	6.2e-25	96.3	1	722-820
1055	P2X_receptor	ATP P2X receptor	9.3e-302	1015.9	1	13-388
1057	zona_pellucida	Zona pellucida-like domain	1.1e-80	281.5	1	268-538
1057	trefoil	Trefoil (P-type) domain	0.02	9.1	1	224-262
1058	Aa_trans	Transmembrane amino acid transporter protein	9.4e-09	42.5	1	30-389
1059	7tm_1	7 transmembrane receptor (rhodopsin family)	2.1e-13	44.8	1	33-280
1060	Na_sulph_symp	Sodium:sulfate symporter transmembrane	1.2e-143	490.7	1	16-554
1064	ERG4_ERG24	Ergosterol biosynthesis ERG4/ERG24 family	1.1e-103	357.8	1	7-350
1066	ig	Immunoglobulin domain	4.7e-16	56.8	3	42-95:135-192:231-288
1067	7tm_1	7 transmembrane receptor (rhodopsin family)	2.3e-33	108.1	1	28-266
1069	pkinase	Protein kinase domain	4.9e-68	239.4	1	298-578
1069	Activin_recip	Activin types I and II receptor domain	1.6e-27	104.8	1	20-107

Table 4A

454

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1070	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.00018	13.2	1	3-177
1071	MHC_I	Class I Histocompatibility antigen, domains	4.4e-14	55.1	1	24-196
1071	ig	Immunoglobulin domain	2.8e-07	28.5	1	218-284
1073	sugar_tr	Sugar (and other) transporter	0.028	-126.7	1	48-528
1074	sugar_tr	Sugar (and other) transporter	0.028	-126.7	1	136-616
1075	ank	Ank repeat	1.7e-45	164.6	6	31-63:64-96:97-129:130-162:163-195:196-228
1076	7tm_1	7 transmembrane receptor (rhodopsin family)	5.6e-12	40.2	1	178-349
1077	sugar_tr	Sugar (and other) transporter	0.0032	-100.1	1	46-470
1079	G_glu_transpept	Gamma-glutamyltranspeptidase	2.8e-05	-144.9	1	122-500
1080	TPR	TPR Domain	2.3e-19	77.8	5	28-61:68-101:108-141:148-181:188-221
1081	ACAT	Sterol O-acyltransferase	1.9e-32	121.3	1	300-406
1082	WD40	WD domain, G-beta repeat	2.3e-16	67.8	7	1015-1050:1059-1097:1115-1151:1158-1194:1203-1240:1246-1281:1293-1329
1083	pkinase	Protein kinase domain	4.9e-57	202.9	2	248-492:537-564
1083	Activin_rec	Activin types I and II receptor domain	3.1e-36	133.8	1	26-127
1084	pkinase	Protein kinase domain	4.9e-57	202.9	2	310-554:599-626
1084	Activin_rec	Activin types I and II receptor domain	3.1e-36	133.8	1	26-127
1086	C2	C2 domain	1.7e-06	35.0	1	233-316
1087	Cache	Cache domain	1.5e-25	96.2	2	557-650:960-985
1088	Glyco_hydro_31	Glycosyl hydrolases family 31	4.9e-268	903.8	2	1-92:114-636
1090	HMG_box	HMG (high mobility group) box	3.8e-32	120.2	1	681-749
1091	serpin	Serpin (serine protease inhibitor)	3.1e-195	662.0	1	315-683
1092	trypsin	Trypsin	0.0044	12.4	1	406-526
1094	Aa_trans	Transmembrane amino acid transporter protein	0.0042	-25.4	1	141-551
1097	ATP-synt_C	ATP synthase subunit C	3.3e-18	73.9	1	72-140
1098	aconitase	Aconitase family (aconitate hydratase)	1.4e-198	651.7	2	162-241:321-744
1098	Aconitase_C	Aconitase C-terminal domain	8.9e-72	251.9	1	872-1043
1103	PAP2	PAP2 superfamily	6.3e-15	63.0	1	89-236
1104	ig	Immunoglobulin domain	1.4e-20	71.3	2	80-148:183-251
1106	TCTP	Translationally controlled tumor protein	3.5e-93	323.0	1	1-166

Table 4A  
455

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1109	efhand	EF hand	1.2e-13	58.8	3	33-61:102-130:138-166
1110	ATP1G1_PLM_MAT8	ATP1G1/PLM/MAT8 family	1.8e-13	58.2	1	92-146
1112	ldl_recept_a	Low-density lipoprotein receptor domain	0.00073	26.3	1	115-153
1112	CUB	CUB domain	0.002	-3.5	1	9-109
1115	7tm_1	7 transmembrane receptor (rhodopsin family)	3.9e-21	69.3	1	139-317
1116	7tm_1	7 transmembrane receptor (rhodopsin family)	1.4e-24	80.3	3	2-28:114-275:348-363
1119	tsp_1	Thrombospondin type 1 domain	4.5e-38	139.9	10	149-198:306-364:571-626:631-696:707-761:841-889:970-1021:1099-1148:1219-1269:1342-1398
1120	HMG14_17	HMG14 and HMG17	1.3e-34	128.4	1	2-86
1125	ig	Immunoglobulin domain	4.7e-20	69.7	3	84-153:185-255:292-347
1127	ig	Immunoglobulin domain	1.6e-10	38.9	1	42-112
1131	C2	C2 domain	5.7e-19	76.5	2	167-257:667-750
1133	7tm_1	7 transmembrane receptor (rhodopsin family)	6.9e-29	93.9	1	104-352
1134	MIP	Major intrinsic protein	3.9e-40	125.5	3	80-189:197-262:308-325
1136	sugar_tr	Sugar (and other) transporter	0.024	-124.9	1	23-504
1137	ATP-synt_C	ATP synthase subunit C	3.3e-35	130.4	2	14-79:90-155
1138	DHDPS	Dihydrodipicolinate synthetase family	4.4e-32	120.0	1	34-325
1139	SSF	Sodium:solute symporter family	1.5e-48	174.7	2	50-461:569-953
1143	7tm_1	7 transmembrane receptor (rhodopsin family)	2.2e-50	162.2	1	78-332
1145	PID	Phosphotyrosine interaction domain (PTB/PID)	1.9e-94	327.2	2	488-627:661-782
1145	WW	WW domain	2.5e-08	41.1	1	373-401
1148	C2	C2 domain	2.1e-30	114.4	2	87-165:240-320
1152	7tm_1	7 transmembrane receptor (rhodopsin family)	9.6e-31	99.8	1	161-410
1153	WD40	WD domain, G-beta repeat	0.013	22.1	4	26-62:71-109:236-271:430-467
1155	aa_permeases	Amino acid permease	5.6e-06	-179.9	1	193-613
1158	dUTPase	dUTPase	5.1e-29	109.8	1	46-167
1158	MIP	Major intrinsic protein	3.8e-28	88.8	1	2-56
1158	rvp	Retroviral aspartyl protease	2.1e-22	85.1	1	179-280
1158	G-patch	G-patch domain	0.00095	25.9	1	285-329
1160	laminin_EGF	Laminin EGF-like (Domains III and V)	9.5e-132	451.1	14	299-353:356-423:426-466:490-532:535-578:584-627:630-680:683-727:1265-

Table 4A

456

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						1308:1311- 1352:1355- 1401:1404- 1452:1686- 1730:1733-1786
1160	laminin_Nterm	Laminin N-terminal (Domain VI)	4e-77	269.6	1	47-297
1160	laminin_G	Laminin G domain	6.1e-31	116.2	3	2625-2748:3015- 3138:3185-3313
1160	laminin_B	Laminin B (Domain IV)	1.1e-09	45.6	1	1517-1651
1161	EGF	EGF-like domain	2.2e-26	101.1	5	29-57:60-88:104- 140:147-178:185- 216
1162	interferon	Interferon alpha/beta domain	1.5e-17	71.7	1	16-171
1163	IL1	Interleukin-1 / 18	8.7e-21	82.5	1	19-159
1164	PDGF	Platelet-derived growth factor (PDGF)	1.8e-51	184.4	1	52-130
1165	IL1	Interleukin-1 / 18	3.7e-23	90.4	1	11-144
1167	Palm_thioest	Palmitoyl protein thioesterase	2.4e-216	732.1	1	28-282
1168	serpin	Serpin (serine protease inhibitor)	1.6e-202	686.2	1	47-415
1170	Defensin_propep	Defensin propeptide	1.2e-26	102.0	1	38-90
1170	defensins	Mammalian defensin	7.3e-14	59.5	1	103-131
1172	Y_phosphatase	Protein-tyrosine phosphatase	1.8e-110	380.4	1	271-499
1179	Metallophos	Calcineurin-like phosphoesterase	1.1e-11	52.3	1	70-285
1184	MHC_I	Class I Histocompatibility antigen, domains	6.2e-06	-5.9	1	29-205
1187	PLA2_B	Lysophospholipase catalytic domain	8.3e-53	188.9	1	357-800
1187	C2	C2 domain	1.6e-06	35.1	1	46-129
1189	lipocalin	Lipocalin / cytosolic fatty-acid binding	2.6e-39	144.0	1	39-188
1189	Kunitz_BPTI	Kunitz/Bovine pancreatic trypsin inhibito	4.7e-23	90.0	2	231-273:274-319
1190	ig	Immunoglobulin domain	4e-07	37.1	1	38-116
1191	ig	Immunoglobulin domain	4e-07	37.1	1	38-116
1200	zf-DHHC	DHHC zinc finger domain	3.7e-15	63.8	1	87-149
1201	KRAB	KRAB box	6.6e-22	86.2	1	33-73
1202	KRAB	KRAB box	6.6e-22	86.2	1	33-73
1204	trypsin	Trypsin	3e-31	117.2	1	56-239
1212	LRR	Leucine Rich Repeat	2.5e-54	193.9	15	73-96:97-122:123- 149:150-175:176- 195:197-220:221- 246:247-270:292- 317:318-337:339- 362:363-388:389- 408:410-433:434- 459
1213	Zn_carbOpept	Zinc carboxypeptidase	4.3e-62	219.7	1	50-288
1221	C1q	C1q domain	3.4e-41	150.3	1	134-258
1221	Collagen	Collagen triple helix repeat	2.6e-10	47.7	1	41-100



Table 4A

457

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		(20 copies)				
1223	pkinese	Protein kinase domain	9e-75	261.8	1	158-417
1223	SH3	SH3 domain	0.019	7.4	1	423-478
1223	UBA	UBA/TS-N domain	0.037	20.6	1	987-1026
1224	vwa	von Willebrand factor type A domain	1.2e-05	29.0	1	51-242
1225	abhydrolase	alpha/beta hydrolase fold	3.7e-12	53.8	1	111-390
1227	lectin_c	Lectin C-type domain	9.6e-29	108.9	1	54-166
1237	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	1.5e-09	-1.3	1	242-385
1257	LRR	Leucine Rich Repeat	9e-65	228.6	14	78-101:102-125:126-149:150-173:174-197:198-221:222-245:246-269:270-293:294-317:318-341:342-365:366-389:390-413
1257	LRRCT	Leucine rich repeat C-terminal domain	0.011	22.3	1	423-475
1263	Pep_M12B_prop ep	Reprolysin family propeptide	1.8e-28	108.1	1	75-191
1263	disintegrin	Disintegrin	2.7e-11	51.0	1	419-494
1263	Reprolysin	Reprolysin (M12B) family zinc metallo	7.3e-07	-22.2	1	206-402
1266	Clusterin	Clusterin	2.3e-298	1004.6	1	2-394
1267	serpin	Serpin (serine protease inhibitor)	3.8e-140	479.0	1	80-425
1268	serpin	Serpin (serine protease inhibitor)	6.8e-162	551.2	1	98-461
1269	COX6C	Cytochrome c oxidase subunit VIc	2.3e-38	140.9	1	1-75
1273	Rhodanese	Rhodanese-like domain	1.5e-48	174.7	2	16-138:165-266
1274	trypsin	Trypsin	2.9e-80	280.1	1	121-346
1279	Calpain_III	Calpain large subunit, domain III	1.5e-25	98.3	1	30-179
1287	zf-C2H2	Zinc finger, C2H2 type	4.4e-66	233.0	9	173-195:201-223:229-251:257-279:285-307:313-335:341-363:369-391:397-419
1287	KRAB	KRAB box	4.9e-23	90.0	1	14-54
1294	vwc	von Willebrand factor type C domain	1.3e-09	45.4	1	69-127
1295	zf-C2H2	Zinc finger, C2H2 type	5e-50	179.6	8	263-285:319-341:347-369:375-397:403-425:431-453:459-481:487-510
1295	KRAB	KRAB box	8.9e-26	99.1	1	8-48
1296	zf-C2H2	Zinc finger, C2H2 type	3.2e-113	389.6	13	190-212:218-240:246-268:274-296:302-324:330-352:358-380:386-408:414-436:442-

Table 4A  
458

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						464:470-492:498-520:526-548
1296	KRAB	KRAB box	6.7e-28	106.1	1	14-54
1299	hemopexin	Hemopexin	7e-25	96.1	4	279-322:324-367:369-415:417-461
1299	Peptidase_M10	Matrixin	8.4e-14	58.8	1	1-176
1305	DUF25	Domain of unknown function DUF25	5.3e-76	265.9	1	40-162
1307	kazal	Kazal-type serine protease inhibitor domain	1.3e-12	55.3	1	21-68
1309	Defensin_propep	Defensin propeptide	1.2e-26	102.0	1	74-126
1309	defensins	Mammalian defensin	7.3e-14	59.5	1	139-167
1312	ras	Ras family	2.7e-59	210.4	1	64-253
1314	zf-C2H2	Zinc finger, C2H2 type	1.1e-51	185.1	8	163-185:191-213:219-241:247-269:275-297:303-325:331-353:359-381
1315	Keratin_B2	Keratin, high sulfur B2 protein	0.0035	-30.0	1	13-154
1316	Keratin_B2	Keratin, high sulfur B2 protein	3.6e-15	63.8	1	23-154
1317	Keratin_B2	Keratin, high sulfur B2 protein	8.2e-20	79.2	2	20-163:164-295
1319	zf-C2H2	Zinc finger, C2H2 type	2.9e-52	187.1	7	33-55:61-83:89-111:117-139:145-167:173-195:201-223
1321	Keratin_B2	Keratin, high sulfur B2 protein	0.00014	-8.3	1	23-193
1331	ig	Immunoglobulin domain	9.7e-06	32.5	1	24-78
1332	Kunitz_BPTI	Kunitz/Bovine pancreatic trypsin inhibitor	2.2e-18	74.5	1	76-126
1333	zf-C2H2	Zinc finger, C2H2 type	4.6e-55	196.3	7	199-221:227-249:255-277:283-305:311-333:339-361:367-389
1333	KRAB	KRAB box	5.8e-24	93.0	1	6-46
1336	vwa	von Willebrand factor type A domain	0.0038	-8.1	1	32-193
1337	vwa	von Willebrand factor type A domain	0.0038	-8.1	1	32-193
1339	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.0088	-16.3	1	4-203
1340	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.0019	-4.7	1	4-194
1343	ig	Immunoglobulin domain	1.5e-13	58.5	2	81-158:204-278
1347	ig	Immunoglobulin domain	1.2e-07	38.9	2	168-226:276-330
1348	ig	Immunoglobulin domain	1.5e-18	75.1	3	157-217:269-325:373-427
1353	LRR	Leucine Rich Repeat	2e-39	144.4	9	120-143:144-167:168-191:192-

Table 4A  
459

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						215:216-239:240-263:264-287:311-334:335-358
1353	LRRNT	Leucine rich repeat N-terminal domain	0.0066	23.1	1	92-119
1356	ig	Immunoglobulin domain	2.1e-17	71.3	3	37-106:138-208:245-300
1357	CD36	CD36 family	3.9e-231	781.3	1	5-445
1358	DUF139	Cysteine rich repeat (DUF139)	1.2e-08	42.2	4	195-211:245-261:288-304:316-332
1359	ig	Immunoglobulin domain	1.6e-06	35.1	2	41-124:156-230
1359	Gag_MA	Matrix protein (MA), p15	0.0044	-25.2	1	291-429
1360	ion_trans	Ion transport protein	0.0031	24.2	1	285-445
1363	spectrin	Spectrin repeat	9.5e-22	85.7	6	17-121:124-226:229-340:372-476:678-785:788-896
1366	C1q	C1q domain	2.1e-31	117.8	1	266-390
1366	Collagen	Collagen triple helix repeat (20 copies)	0.00029	19.8	1	182-241
1367	rnaseA	Pancreatic ribonuclease	1.1e-34	128.7	1	27-132
1368	3HCDH_N	3-hydroxyacyl-CoA dehydrogenase, NAD binding	0.0021	-60.2	1	50-141
1371	voltage_CLC	Voltage gated chloride channel	3.1e-199	675.3	1	92-528
1371	CBS	CBS domain	1.2e-24	95.4	2	559-617:761-814
1375	Peptidase_C1	Papain family cysteine protease	1e-120	414.5	1	114-332
1376	7tm_1	7 transmembrane receptor (rhodopsin family)	2.6e-50	180.6	1	48-454
1378	7tm_2	7 transmembrane receptor (Secretin family)	4.9e-09	32.5	1	250-500
1380	MAGE	MAGE family	1.4e-91	317.7	1	3-231
1381	ion_trans	Ion transport protein	0.066	16.8	1	158-323
1382	ig	Immunoglobulin domain	7e-12	52.9	2	37-128:160-241
1385	MHC_I	Class I Histocompatibility antigen, domains	9.3e-06	-8.4	1	29-204
1388	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-tr	7e-88	305.4	1	45-184
1390	7tm_1	7 transmembrane receptor (rhodopsin family)	5.3e-06	21.2	1	1-210
1393	7tm_1	7 transmembrane receptor (rhodopsin family)	1.6e-12	55.0	1	99-329
1397	fn3	Fibronectin type III domain	4.7e-78	272.7	5	533-619:631-717:734-833:854-939:951-1039
1397	ig	Immunoglobulin domain	3.7e-42	153.5	5	154-227:261-318:362-418:453-511:1253-1316
1404	ig	Immunoglobulin domain	2.9e-08	40.9	2	42-99:139-198
1409	ig	Immunoglobulin domain	8.1e-30	112.5	3	142-199:263-321:359-438
1414	mito_carr	Mitochondrial carrier	3.5e-64	226.7	3	40-130:137-

Table 4A  
460

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		protein				227:238-322
1416	ig	Immunoglobulin domain	0.0011	25.6	1	57-112
1417	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	6.6e-06	33.1	1	27-67
1418	Synaptophysin	Synaptophysin / synaptoporin	9.3e-73	255.1	1	27-273
1419	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	5.2e-18	73.3	1	4-165
1421	Ammonium_transp	Ammonium Transporter Family	8.5e-18	72.6	1	25-383
1423	rnaseA	Pancreatic ribonuclease	4.1e-29	110.1	1	32-136
1427	LRR	Leucine Rich Repeat	4.5e-26	100.0	6	3-26:27-50:51-74:76-99:100-123:124-144
1427	fn3	Fibronectin type III domain	2.7e-08	41.0	1	334-417
1427	ig	Immunoglobulin domain	2.4e-07	37.9	1	225-283
1427	LRRCT	Leucine rich repeat C-terminal domain	7.2e-06	32.9	1	164-209
1428	GRAM	GRAM domain	0.067	13.1	1	177-245
1435	Rhomboid	Rhomboid family	1.5e-45	164.8	1	169-326
1437	Sema	Sema domain	7e-202	684.1	1	102-513
1437	tsp_1	Thrombospondin type 1 domain	1.1e-42	155.3	6	589-637:644-695:702-746:833-883:890-940:945-985
1437	PSI	Plexin repeat	2.5e-14	61.0	1	531-578
1438	ank	Ankyrin repeat	2e-15	64.7	3	336-368:383-418:468-498
1438	ion_trans	Ion transport protein	9.4e-07	35.9	1	615-817
1439	ank	Ankyrin repeat	3.3e-12	54.0	3	417-449:464-499:549-579
1439	ion_trans	Ion transport protein	9.4e-07	35.9	1	696-898
1442	LRR	Leucine Rich Repeat	3.2e-20	80.6	5	57-80:81-104:105-128:129-152:153-176
1442	LRRCT	Leucine rich repeat C-terminal domain	2.3e-07	37.9	1	186-240
1442	LRRNT	Leucine rich repeat N-terminal domain	0.00013	28.8	1	27-55
1443	7tm_2	7 transmembrane receptor (Secretin family)	7.2e-23	89.4	1	166-476
1443	HMG_box	HMG (high mobility group) box	3.1e-05	30.8	1	3-76
1443	GPS	Latrophilin/CL-1-like GPS domain	0.0066	19.1	1	110-162
1444	DUF214	Predicted permease	0.0045	-15.6	1	136-276
1452	mito_carr	Mitochondrial carrier protein	8.4e-58	205.5	2	27-123:124-217
1454	mito_carr	Mitochondrial carrier protein	0.00049	22.3	2	79-179:198-287
1455	WD40	WD domain, G-beta repeat	2.7e-26	100.8	4	64-99:105-141:235-273:279-315
1456	Furin-like	Furin-like cysteine rich	6.1e-89	308.9	1	172-328

Table 4A

461

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		region				
1456	Recep_L_domain	Receptor L domain	3.1e-80	280.0	2	46-169:345-471
1456	pkinese	Protein kinase domain	7e-65	229.0	1	986-1258
1462	Peptidase_M10	Matrixin	1.5e-79	277.7	1	62-227
1462	hemopexin	Hemopexin	1.4e-10	48.6	1	309-350
1463	Armadillo_seg	Armadillo/beta-catenin-like repeat	0.028	21.0	4	549-591:641-685:686-737:738-783
1466	C2	C2 domain	6.5e-07	36.4	1	219-304
1468	RTC	RNA 3'-terminal phosphate cyclase	0.0006	-	1	3-221
1476	CN_hydrolase	Carbon-nitrogen hydrolase	0.0014	-83.3	1	64-341
1477	zf-C2H2	Zinc finger, C2H2 type	4.2e-06	33.7	3	267-289:294-317:419-442
1478	Euk_porin	Eukaryotic porin	3.3e-06	-42.1	1	1-142
1484	BTB	BTB/POZ domain	1.8e-09	44.9	1	52-194
1487	rvt	Reverse transcriptase	5.6e-56	199.4	1	241-507
1489	Galactosyl_T_2	Galactosyltransferase	3.6e-175	595.3	1	97-404
1491	GBP_C	Guanylate-binding protein, C-terminal domain	1.8e-32	117.0	1	23-125
1493	NUDIX	MutT-like domain	0.0018	23.5	1	297-439
1496	Big_2	Bacterial Ig-like domain (group 2)	4.5e-16	66.8	1	1177-1258
1497	ig	Immunoglobulin domain	9.9e-13	46.1	1	38-112
1498	aldo_ket_red	Aldo/keto reductase family	6.8e-37	126.5	2	207-237:245-359
1502	Caveolin	Caveolin	1.6e-06	13.3	1	90-262
1504	sushi	Sushi domain (SCR repeat)	8.2e-16	66.0	2	223-281:286-344
1511	zf-C2H2	Zinc finger, C2H2 type	1.1e-16	68.9	2	341-363:369-391
1511	KRAB	KRAB box	6e-05	22.1	1	231-275
1516	CBS	CBS domain	2.8e-33	124.0	3	189-243:264-317:336-389
1517	zf-C2H2	Zinc finger, C2H2 type	2.1e-128	440.0	18	279-301:307-329:335-357:363-385:391-413:419-441:447-469:475-497:503-525:868-890:896-918:924-946:952-974:980-1002:1008-1030:1036-1058:1064-1086:1092-1114
1517	KRAB	KRAB box	5.3e-69	242.7	2	74-136:684-735
1518	ig	Immunoglobulin domain	1.1e-07	29.8	1	38-102
1521	CoA_trans	Coenzyme A transferase	1.1e-161	550.5	2	42-261:271-472
1521	TGF-beta	Transforming growth factor beta like domain	6.8e-64	225.7	1	637-741
1523	Reprolysin	Reprolysin (M12B) family zinc metallo	4.4e-95	329.3	1	223-422
1523	Pep_M12B_prop cp	Reprolysin family propeptide	6.7e-30	112.8	1	121-234
1523	disintegrin	Disintegrin	1.1e-18	59.5	1	437-513
1523	EGF	EGF-like domain	0.0092	22.6	1	661-693
1524	DSPc	Dual specificity phosphatase, catalytic doma	3.6e-71	249.9	1	102-241

Table 4A  
462

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1525	7tm_1	7 transmembrane receptor (rhodopsin family)	5.6e-14	46.6	1	1-53
1527	Spermine_synth	Spermine/spermidine synthase	6e-13	56.4	1	254-441
1528	efhand	EF hand	3.8e-10	47.1	4	65-91:145-173:329-357:365-393
1530	rvt	Reverse transcriptase	6e-68	239.2	1	125-395
1531	RPH3A_effector	Rabphilin-3A effector domain	1.9e-09	-6.1	1	161-343
1533	TPR	TPR Domain	4.1e-10	47.0	3	220-253:444-477:478-511
1534	SAM	SAM domain (Sterile alpha motif)	0.046	18.0	1	688-752
1538	pkinase	Protein kinase domain	7.2e-42	152.5	2	42-76:82-242
1540	zf-C2H2	Zinc finger, C2H2 type	0.00035	27.3	1	725-750
1541	COLFI	Fibrillar collagen C-terminal domain	8.5e-216	553.4	1	624-841
1542	EGF	EGF-like domain	2.6e-11	51.0	3	916-947:1004-1044:1050-1094
1543	ig	Immunoglobulin domain	1.3e-12	45.7	1	38-112
1544	UCH-1	Ubiquitin carboxyl-terminal hydrolases famil	9.1e-15	62.5	1	477-508
1544	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	6.4e-13	56.3	1	759-811
1544	zf-UBP	Zn-finger in ubiquitin-hydrolases and other	0.00069	18.9	1	350-420
1546	LRR	Leucine Rich Repeat	2.4e-30	114.3	9	132-155:156-179:180-199:201-224:250-269:271-294:295-318:319-341:342-365
1546	LRRNT	Leucine rich repeat N-terminal domain	0.00082	26.1	1	62-84
1551	Peptidase_S9	Prolyl oligopeptidase family	4.3e-10	36.8	1	397-469
1552	Orn_DAP_Arg_deC	Pyridoxal-dependent decarboxylase, C-	1.5e-14	55.5	1	150-202
1552	Orn_Arg_deC_N	Pyridoxal-dependent decarboxylase, py	4.4e-11	40.3	1	19-117
1555	rvt	Reverse transcriptase	5e-17	67.6	1	616-718
1557	rvt	Reverse transcriptase	1.1e-68	241.6	1	247-520
1558	RhoGEF	RhoGEF domain	1.5e-25	98.4	1	971-1155
1558	PDZ	PDZ domain (Also known as DHR or GLGF)	4.4e-12	53.6	1	159-234
1558	PH	PH domain	0.0017	17.5	1	1199-1312
1559	pkinase	Protein kinase domain	1.9e-24	92.1	2	151-231:309-419
1560	FH2	Formin Homology 2 Domain	4.4e-111	382.4	1	595-1038
1565	HMG_CoA_synt	Hydroxymethylglutaryl-coenzyme A synthas	6.5e-300	1009.7	1	50-582
1568	GCV_H	Glycine cleavage H-protein	1.6e-76	267.6	1	65-185
1577	Peptidase_M1	Peptidase family M1	1.3e-114	332.5	3	42-161:225-354:365-399
1579	rrm	RNA recognition motif.	0.0068	23.0	1	823-891
1580	serpin	Serpin (serine protease	8.9e-135	458.1	3	37-241:352-

Table 4A  
463

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		inhibitor)				475:504-533
1581	CH	Calponin homology (CH) domain	0.012	13.5	1	84-181
1582	LRR	Leucine Rich Repeat	2.2e-37	137.6	12	70-93:94-117:118-141:142-165:166-189:190-213:214-237:238-261:262-285:286-310:311-335:336-359
1582	LRRCT	Leucine rich repeat C-terminal domain	8.1e-12	52.7	1	369-421
1582	ig	Immunoglobulin domain	1.9e-08	32.2	1	438-499
1582	LRRNT	Leucine rich repeat N-terminal domain	0.0099	22.5	1	28-68
1587	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.00027	18.1	1	168-213
1587	IBR	IBR domain	0.0071	10.2	1	237-298
1588	rvt	Reverse transcriptase	1.2e-05	25.4	1	192-250
1589	rvt	Reverse transcriptase	1e-10	44.2	1	24-95
1591	TPR	TPR Domain	3.8e-13	57.1	3	254-287:288-321:322-355
1592	PHD	PHD-finger	0.042	2.9	2	684-725:1117-1163
1595	Peptidase_M1	Peptidase family M1	2e-113	329.1	1	100-279
1597	IPPe	Inositol polyphosphate phosphatase family, c	1.1e-15	57.7	1	50-112
1600	thyroglobulin_1	Thyroglobulin type-1 repeat	3.1e-06	34.1	1	38-118
1604	ig	Immunoglobulin domain	0.014	13.3	1	34-59
1606	zf-CXXC	CXXC zinc finger	1.7e-21	84.9	1	50-96
1606	F-box	F-box domain	7.2e-06	32.9	2	545-602:715-762
1606	PHD	PHD-finger	0.062	1.3	1	207-271
1613	rrm	RNA recognition motif.	2.9e-14	60.8	1	151-221
1614	SRCR	Scavenger receptor cysteine-rich domain	2.2e-26	101.1	1	52-145
1617	B56	Protein phosphatase 2A regulatory B subunit	6.1e-35	124.7	1	146-218
1618	S1	S1 RNA binding domain	1.2e-10	44.3	1	352-429
1619	trypsin	Trypsin	5.1e-39	125.5	1	30-122
1620	STphosphatase	Ser/Thr protein phosphatase	1.7e-71	245.5	1	66-236
1623	FlgI	Flagellar P-ring protein	5.4e-81	282.5	1	79-231
1623	FlgH	Flagellar L-ring protein	7.5e-34	110.4	1	21-73
1625	Semialdehyde_dh C	Semialdehyde dehydrogenase, dimerisat	8.1e-61	215.5	1	137-308
1625	Semialdehyde_dh	Semialdehyde dehydrogenase, NAD bindi	2.6e-34	109.5	1	20-128
1628	HTH_1	Bacterial regulatory helix-turn-helix protei	2.6e-33	124.1	1	7-146
1631	PEP-utilizers_C	PEP-utilizing enzyme, TIM barrel doma	9.1e-64	225.3	1	53-163
1632	ACR_tran	AcrB/AcrD/AcrF family	3.1e-15	58.8	1	188-254
1632	OTCace	Aspartate/ornithine carbamoyltransferase, A	1.7e-09	35.8	1	459-514
1632	OTCace_N	Aspartate/ornithine carbamoyltransferase, c	0.44	8.5	1	411-426
1632	Replicase	Replicase family	2.9	-	1	399-576

Table 4A

464

SEQ ID	Model	Description	E-value	Score	Repeats	Position
				102.2		
1634	ribonuclease_T2	Ribonuclease T2 family	4.3e-17	70.2	1	78-138
1638	efhand	EF hand	5.6e-10	46.6	4	38-66:74-102:111-139:147-175
1638	RrnaAD	Ribosomal RNA adenine dimethylase	5	2.6	1	133-173
1643	Ribosomal_L23	Ribosomal protein L23	1.2e-14	53.9	1	197-276
1645	tsp_1	Thrombospondin type 1 domain	0.0026	21.0	5	420-476:480-534:541-601:607-663:669-718
1647	Arginosuc_synth	Arginosuccinate synthase	2.8e-68	218.8	2	48-126:128-189
1651	vwa	von Willebrand factor type A domain	1.1e-33	125.4	1	89-326
1652	ig	Immunoglobulin domain	6.1e-10	37.1	2	34-117:220-293
1654	PP2C	Protein phosphatase 2C	3.4e-05	21.8	1	86-138
1658	thiore	Thioredoxin	1.1e-105	354.6	2	324-432:459-570
1662	Hist_deacetyl	Histone deacetylase family	5.5e-85	295.7	2	857-924:1247-1434
1666	wap	WAP-type (Whey Acidic Protein) 'four-disulf	1.5e-21	85.1	1	72-116
1667	KRAB	KRAB box	4.3e-43	156.6	1	106-168
1672	ig	Immunoglobulin domain	1e-13	49.3	3	146-203:245-295:331-402
1674	Transposase_22	L1 transposable element	1.5e-51	184.7	1	6-281
1676	Ribosomal_S2	Ribosomal protein S2	9.5e-09	34.8	1	51-134
1678	ig	Immunoglobulin domain	0.036	12.0	1	35-78
1684	adh_short	short chain dehydrogenase	2.3e-32	120.9	1	2-244
1686	pkinase	Protein kinase domain	2.5e-50	180.6	2	553-770:875-896
1687	HRM	Hormone receptor domain	7.8e-05	29.5	1	82-134
1688	SH3	SH3 domain	1.9e-21	84.7	1	150-206
1688	SH2	SH2 domain	1.2e-08	28.6	1	214-233
1688	RGS	Regulator of G protein signaling domain	0.0003	19.5	1	75-100
1693	LRR	Leucine Rich Repeat	5.3e-34	126.4	10	505-528:529-554:555-573:575-598:613-632:634-657:659-684:686-708:709-732:738-763
1694	SH3	SH3 domain	0.00093	18.2	1	19-73
1701	cystatin	Cystatin domain	8.9e-11	41.4	1	32-78
1705	Ribosomal_S26e	Ribosomal protein S26e	2.1e-16	68.0	1	133-231
1706	ank	Ank repeat	4.9e-30	113.2	3	381-413:414-446:450-482
1706	TPR	TPR Domain	1.4e-06	35.3	4	3-36:43-76:164-197:205-238
1706	LRR	Leucine Rich Repeat	0.0026	24.4	4	716-743:744-764:775-802:835-859
1708	zf-C2H2	Zinc finger, C2H2 type	3.6e-23	90.4	4	51-74:205-227:233-255:261-284
1709	Pep_M12B_prop ep	Reprolysins family propeptide	0.048	-17.2	1	120-196
1710	TIG	IPT/TIG domain	1.4e-101	350.8	10	620-705:708-



Table 4A  
465

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						788:793-876:883-988:1084-1169:1174-1254:1256-1336:1341-1423:1424-1511:1516-1602
1711	HK	Hydroxyethylthiazole kinase family	5.4e-09	35.0	1	12-39
1711	pp-binding	Phosphopantetheine attachment site	5.4e-09	38.4	1	220-255
1712	IGFBP	Insulin-like growth factor binding pr	2.1e-26	97.4	1	40-99
1712	thyroglobulin_1	Thyroglobulin type-1 repeat	4.6e-05	29.5	1	264-323
1715	rvt	Reverse transcriptase	7.8e-26	99.3	1	18-147
1717	A2M	Alpha-2-macroglobulin family	0	1259.4	4	776-879:880-1111:1112-1172:1243-1530
1717	A2M_N	Alpha-2-macroglobulin family N-terminal regi	0	1092.6	1	75-684
1717	TCTP	Translationally controlled tumor protein	0.0013	-47.0	1	1172-1249
1718	cadherin	Cadherin domain	2e-29	111.2	3	68-159:334-425:439-543
1722	p450	Cytochrome P450	6.3e-48	172.0	1	270-506
1725	ig	Immunoglobulin domain	0.0079	14.2	1	34-116
1726	ig	Immunoglobulin domain	4.2e-05	21.5	2	87-131:145-163
1727	Synaptophysin	Synaptophysin / synaptoporin	2.2e-229	775.4	1	12-307
1729	PA	PA domain	2.4e-29	110.9	1	521-626
1729	Glyco_hydro_47	Glycosyl hydrolase family 47	6.2e-13	28.4	1	99-364
1731	7tm_1	7 transmembrane receptor (rhodopsin family)	1.1e-65	210.8	2	50-117:118-293
1732	ig	Immunoglobulin domain	1.7e-18	64.6	2	49-125:169-239
1734	p450	Cytochrome P450	1e-16	61.8	3	50-106:162-228:298-425
1735	profilin	Profilin	1.5e-15	60.5	1	966-1071
1735	Transposase_22	L1 transposable element	0.00026	-102.7	1	731-1016
1742	PAS	PAS domain	3.2e-07	31.9	3	360-413:474-507:581-613
1744	Dynein_heavy	Dynein heavy chain	9.8e-251	846.4	1	34-781
1745	SCP	SCP-like extracellular protein	9.2e-36	132.2	1	4-180
1749	lipocalin	Lipocalin / cytosolic fatty-acid binding pr	6.5e-05	21.0	1	42-76
1751	SDF	Sodium:dicarboxylate symporter family	1.8e-47	171.2	2	1-34:315-396
1753	ig	Immunoglobulin domain	9.4e-19	65.5	5	49-124:162-219:261-310:345-400:432-496
1753	fn3	Fibronectin type III domain	0.00052	26.8	1	514-606
1757	IPPe	Inositol polyphosphate phosphatase family, c	3e-68	240.1	3	64-350:363-621:753-796

Table 4A

466

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1762	SSF	Sodium:solute symporter family	4.2e-256	864.2	1	39-477
1763	LRR	Leucine Rich Repeat	3.4e-37	137.0	11	61-84:85-108:109-132:133-156:157-180:181-204:205-228:253-276:277-300:301-324:349-371
1763	ig	Immunoglobulin domain	9.1e-09	33.3	1	428-486
1763	LRRCT	Leucine rich repeat C-terminal domain	3.1e-07	37.5	1	358-411
1763	LRRNT	Leucine rich repeat N-terminal domain	0.00023	27.9	1	30-59
1764	Reprolysin	Reprolysin (M12B) family zinc metalloprote	1.1e-21	85.5	1	156-359
1764	tsp_1	Thrombospondin type 1 domain	2.5e-06	34.5	1	422-491
1768	ANF_receptor	Receptor family ligand binding region	1.9e-34	125.9	1	68-174
1769	C2	C2 domain	2.5e-66	233.8	2	249-337:383-471
1771	cadherin	Cadherin domain	2e-80	280.6	5	140-235:249-343:359-451:465-556:607-697
1775	7tm_1	7 transmembrane receptor (rhodopsin family)	3.2e-34	110.9	2	34-130:147-330
1778	synaptobrevin	Synaptobrevin	2.3e-15	64.5	1	104-192
1780	MHC_I	Class I Histocompatibility antigen, domains	8.1e-93	321.8	1	25-178
1780	ig	Immunoglobulin domain	9.9e-08	30.0	1	194-259
1782	ABC_tran	ABC transporter	7.2e-19	76.1	1	448-626
1783	SNF	Sodium:neurotransmitter symporter family	0	1182.5	1	203-747
1784	TIG	IPT/TIG domain	4.6e-33	123.3	2	1032-1119:1121-1206
1784	Plexin_repeat	Plexin repeat	4.1e-08	40.4	3	653-706:806-858:953-1009
1784	Sema	Sema domain	2.8e-06	26.6	2	59-160:576-602
1785	CTF_NFI	CTF/NF-I family	1.6e-167	570.0	1	651-824
1786	TIR	TIR domain	1e-38	132.4	1	791-931
1786	LRR	Leucine Rich Repeat	4.4e-37	136.6	15	143-166:192-215:216-241:242-266:267-290:385-408:409-432:433-456:457-480:481-502:545-569:570-598:599-622:623-643:645-668
1786	LRRCT	Leucine rich repeat C-terminal domain	7.3e-13	56.2	1	675-726
1788	RPEL	RPEL repeat	1.2e-17	72.1	2	222-247:266-291
1789	ACR_tran	AcrB/AcrD/AcrF family	1.8e-191	649.5	1	1-452
1792	Sulfatase	Sulfatase	0.00013	-44.3	1	502-912
1792	3_5_exonuclease	3'-5' exonuclease	0.0041	-14.7	1	764-923
1794	DUF81	Domain of unknown function DUF81	3.2e-49	176.1	1	8-166

Table 4A

467

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1796	LRR	Leucine Rich Repeat	7.6e-09	42.8	2	314-337:338-361
1796	LRRCT	Leucine rich repeat C-terminal domain	1.9e-08	41.5	1	371-421
1796	GPS	Latrophilin/CL-1-like GPS domain	0.0041	20.8	1	887-939
1796	HRM	Hormone receptor domain	0.0076	16.3	1	531-600
1796	7tm_2	7 transmembrane receptor (Secretin family)	0.018	- 101.1	1	935-1252
1798	secY	eubacterial secY protein	9.6e-62	218.5	1	1-287
1801	PTR2	POT family	3.5e-38	130.8	2	4-72:364-505
1802	asp	Eukaryotic aspartyl protease	5.1e-13	41.6	1	133-471
1802	Dak1	Dak1 domain	6.7e-06	- 132.2	1	695-1003
1803	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	4.5e-21	83.4	1	4-181
1804	HCO3_cotransp	HCO3- transporter family	1.3e-69	244.7	3	39-52:108-168:171-228
1805	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	5.1e-17	70.0	1	1278-1339
1805	UCH-1	Ubiquitin carboxyl-terminal hydrolases famil	9.3e-08	39.2	1	785-820
1805	zf-UBP	Zn-finger in ubiquitin-hydrolases and other	0.0049	15.8	1	688-727
1806	Nucleoside_tran	Nucleoside transporter	5.7e-42	152.8	1	127-446
1810	sugar_tr	Sugar (and other) transporter	0.044	- 132.3	1	637-1048
1813	zf-C2H2	Zinc finger, C2H2 type	1e-38	142.1	7	270-292:298-320:414-437:440-463:485-507:513-535:771-793
1815	trypsin	Trypsin	4.6e-56	179.2	1	28-341
1817	TauD	Taurine catabolism dioxygenase TauD, TfdA fa	1.7e-196	666.2	1	107-385
1818	BPD_transp	Binding-protein-dependent transport system	0.0048	23.3	1	129-194
1819	Y_phosphatase	Protein-tyrosine phosphatase	1.3e-103	357.6	1	1277-1494
1819	fn3	Fibronectin type III domain	1.2e-51	185.1	7	461-541:552-634:641-719:730-812:819-900:908-990:998-1087

Table 4B  
468

SEQ ID	Model	Description	E-value	Score	Repeats	Position
912	MIF	Macrophage migration inhibitory factor (MIF)	2.4e-54	194.0	1	124-227
912	GST_C	Glutathione S-transferase, C-terminal domain	7.1e-21	82.8	1	410-525
912	GST_N	Glutathione S-transferase, N-terminal domain	2e-15	64.7	1	324-400
913	tsp_1	Thrombospondin type 1 domain	2.9e-23	90.7	2	124-174:181-231
913	EGF	EGF-like domain	1.6e-20	81.6	7	474-508:514-553:559-591:597-633:692-720:802-836:842-882
913	Keratin_B2	Keratin, high sulfur B2 protein	3	-75.8	1	523-649
913	granulin	Granulin	3.2	-15.2	1	484-510
913	TIL	Trypsin Inhibitor like cysteine rich domai	6.4	-13.2	2	540-597:785-842
913	cadherin	Cadherin domain	6.5	-10.4	1	903-995
914	MIF	Macrophage migration inhibitory factor (MIF)	4.5e-50	179.8	1	2-101
915	MIF	Macrophage migration inhibitory factor (MIF)	6.1e-68	239.1	1	2-129
916	trypsin	Trypsin	2.3e-91	316.9	1	445-675
916	kringle	Kringle domain	4.2e-45	163.3	1	286-367
916	fn2	Fibronectin type II domain	4.1e-25	96.9	1	108-148
916	EGF	EGF-like domain	6.2e-15	63.0	2	164-197:245-278
916	fn1	Fibronectin type I domain	4.8e-10	46.8	1	202-237
916	Keratin_B2	Keratin, high sulfur B2 protein	5.3	-79.6	1	104-211
916	DUF32	Domain of unknown function DUF32	9.6	-154.4	1	445-556
917	laminin_EGF	Laminin EGF-like (Domains III and V)	9.8e-159	540.7	12	282-346:349-409:412-469:472-520:523-565:786-831:834-877:880-925:928-984:987-1036:1039-1094:1095-1139
917	laminin_Nterm	Laminin N-terminal (Domain VI)	5.9e-105	362.1	1	45-280
917	DUF26	Domain of unknown function DUF26	3.2	-13.0	1	1110-1155
917	DUF232	Putative transcriptional regulator	3.2	-24.4	1	1149-1281
917	LIM	LIM domain	3.8	-16.6	1	1125-1189
917	zf-CXXC	CXXC zinc finger	4.5	-8.4	1	367-414
917	Methyltransf D12	D12 class N6 adenine-specific DNA met	5.3	-88.8	1	1404-1654
917	VSP	Giardia variant-specific surface prot	8.4	-265.1	1	864-1168
917	EB	EB module	9.5	-12.4	1	1069-1125
918	EGF	EGF-like domain	3.2e-53	190.2	9	33-68:74-110:116-151:161-197:208-243:247-282:288-323:329-362:368-

Table 4B  
469

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						404
918	CUB	CUB domain	2.9e-17	70.8	1	708-817
918	TIL	Trypsin Inhibitor like cysteine rich domai	1.3	-5.0	1	219-288
918	granulin	Granulin	2.4	-14.4	1	242-284
918	wap	WAP-type (Whey Acidic Protein) 'four-disul	4.7	-8.7	1	6-71
918	Keratin_B2	Keratin, high sulfur B2 protein	6.6	-81.2	1	95-247
918	VSP	Giardia variant-specific surface protein	7.7	-264.1	1	95-403
919	Parathyroid	Parathyroid hormone family	3.1e-82	286.6	1	46-175
920	EGF	EGF-like domain	2.2e-22	87.8	9	2-36:47-82:88-122:128-163:232-263:276-306:319-349:353-394:407-439
921	IL8	Small cytokines (intecrine/chemokine), inter	0.19	-3.2	1	25-65
922	Zn_carbOpept	Zinc carboxypeptidase	8.7e-120	411.4	1	149-429
922	Propep_M14	Carboxypeptidase activation peptide	5.6e-20	79.8	1	7-84
923	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	9.5e-59	208.6	1	62-263
927	FKBP	FKBP-type peptidyl-prolyl cis-trans isomeras	2.4e-41	150.8	1	40-122
930	FAD-oxidase_C	FAD linked oxidases, C-terminal domain	4.1e-88	306.2	1	395-647
930	FAD_binding_4	FAD binding domain	5.2e-78	272.6	2	24-172:219-381
931	UPF0136	Uncharacterised protein family (UPF0136)	0.00031	17.6	1	12-75
933	Fringe	Fringe-like	1.4e-06	-12.8	1	262-487
933	Galactosyl_T	Galactosyltransferase	2.6	-84.8	1	282-439
937	14-3-3	14-3-3 protein	9.5e-124	424.5	1	50-282
938	S_100	S-100/ICaBP type calcium binding domain	3.6e-23	90.4	1	65-108
938	efhand	EF hand	0.0012	25.6	1	114-142
940	ig	Immunoglobulin domain	2.9e-35	130.6	5	38-105:139-206:241-293:326-377:412-487
941	LRR	Leucine Rich Repeat	2e-30	114.5	8	42-63:64-87:88-112:113-136:137-160:161-184:185-208:209-232
941	LRRCT	Leucine rich repeat C-terminal domain	3.6e-07	37.2	1	242-292
941	LRRNT	Leucine rich repeat N-terminal domain	0.46	15.7	1	11-40
944	serpin	Serpin (serine protease inhibitor)	1.6e-133	457.0	1	173-532
947	RUN	RUN domain	8e-44	159.0	1	31-163
947	FYVE	FYVE zinc finger	3.5e-13	57.2	1	529-598
947	Tektin	Tektin family	1.9	-229.2	1	154-473

Table 4B  
470

SEQ ID	Model	Description	E-value	Score	Repeats	Position
947	filament	Intermediate filament protein	4.7	-202.1	1	216-501
947	PAPS_reduct	Phosphoadenosine phosphosulfate reductase	4.9	-69.1	1	305-467
947	metalthio	Metallothionein	7.3	-11.4	1	576-635
950	lectin_c	Lectin C-type domain	2.4e-08	41.1	1	130-234
951	lectin_c	Lectin C-type domain	8	-36.4	1	215-281
951	FAD-oxidase_C	FAD linked oxidases, C-terminal domain	9.7	-84.1	1	60-267
952	Glyco_transf_8	Glycosyl transferase family 8	0.01	-47.7	1	1255-1526
952	Glyco_hydro_20	Glycosyl hydrolase family 20, catalyti	9.8	-224.0	1	391-674
953	TPR	TPR Domain	0.0042	23.8	1	171-204
954	DENN	DENN (AEX-3) domain	4.2	-67.5	1	16-142
956	Branch	Core-2/I-Branching enzyme	2.2e-77	270.5	1	46-310
957	fn3	Fibronectin type III domain	1.1e-15	65.6	1	30-108
958	Peptidase_M10	Matrixin	5.9e-121	415.2	1	58-225
958	hemopexin	Hemopexin	2e-58	207.5	4	305-347:349-393:398-445:447-487
958	Astacin	Astacin (Peptidase family M12A)	0.28	-108.1	1	128-285
958	PG_binding_1	Putative peptidoglycan binding domain	2.4	-9.5	1	44-112
960	Robl_LC7	Roadblock/LC7 domain	1.8e-30	114.7	1	56-147
962	chromo	'chromo' (CHRromatin Organization MOdifier)	2.7e-05	31.1	1	47-84
962	SET	SET domain	1.2	-48.2	1	243-307
964	ig	Immunoglobulin domain	0.0047	23.6	1	34-115
967	7tm_1	7 transmembrane receptor (rhodopsin family)	3.5e-13	57.2	1	41-237
968	adh_zinc	Zinc-binding dehydrogenase	2.4e-61	217.2	1	174-495
968	DUF142	Domain of unknown function DUF142	6.6	-67.7	1	248-332
970	E1_dehydrog	Dehydrogenase E1 component	2.1e-183	622.7	1	129-425
970	transketolase	Transketolase, thiamine diphosphate bin	2.6	-227.4	1	173-410
971	ig	Immunoglobulin domain	5.8e-09	43.2	1	45-122
973	zf-MIZ	MIZ zinc finger	3.2e-32	120.5	1	279-331
973	SAP	SAP domain	2.4e-05	31.2	1	2-36
973	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	5.7	-6.3	1	283-325
975	WD40	WD domain, G-beta repeat	4.3e-09	43.6	3	234-270:446-483:491-528
977	An_peroxidas e	Animal haem peroxidase	6.9e-137	468.2	1	137-701
978	Defensin_pro pep	Defensin propeptide	3.8e-19	77.0	1	32-82
978	defensins	Mammalian defensin	0.0015	25.3	1	95-123
981	ig	Immunoglobulin domain	0.00022	28.0	1	49-130
982	ig	Immunoglobulin domain	5.8e-07	36.6	1	36-112
983	Acetyltransf	Acetyltransferase (GNAT)	0.22	15.5	1	260-334

Table 4B  
471

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		family				
985	tsp_1	Thrombospondin type 1 domain	0.0075	17.1	1	151-206
985	Furin-like	Furin-like cysteine rich region	5.1	-94.5	1	37-149
985	TNFR_c6	TNFR/NGFR cysteine-rich region	9.6	-0.7	1	114-148
987	LRR	Leucine Rich Repeat	5.5e-59	209.4	12	78-101:102-125:126-149:150-173:174-197:198-221:222-245:246-269:270-293:294-317:318-341:342-365
987	LRRNT	Leucine rich repeat N-terminal domain	0.23	18.0	1	24-52
987	Keratin_B2	Keratin, high sulfur B2 protein	6.1	-80.6	1	334-442
988	ParA	ParA family ATPase	0.029	0.0	1	177-283
988	APS_kinase	Adenylylsulfate kinase	0.92	-112.5	1	102-226
988	ArsA_ATPase	Anion-transporting ATPase	0.97	-180.2	1	103-336
989	vwd	von Willebrand factor type D domain	1.1e-38	141.9	1	364-514
989	vwc	von Willebrand factor type C domain	5.3e-35	129.7	5	50-105:108-163:166-224:238-289:301-357
989	TIL	Trypsin Inhibitor like cysteine rich	8.9e-06	32.6	1	629-682
989	Keratin_B2	Keratin, high sulfur B2 protein	1	-68.4	1	264-398
989	TILa	TILa domain	1.4	-2.7	1	301-353
989	fn1	Fibronectin type I domain	2.3	4.7	1	166-198
989	Metallothio_P EC	Plant PEC family metallothionein	4.8	-38.2	1	587-651
990	Cys_knot	Cystine-knot domain	3.2	-31.3	1	227-313
992	ig	Immunoglobulin domain	2.9e-13	57.5	2	34-111:150-220
994	ig	Immunoglobulin domain	1.2e-26	102.0	3	43-110:143-197:230-291
995	SCP	SCP-like extracellular protein	5e-41	149.7	1	4-173
995	granulin	Granulin	6.9	-17.5	1	170-199
996	ig	Immunoglobulin domain	0.014	22.0	1	153-206
999	RNase_PH	3' exoribonuclease family, domain 1	6.1e-23	89.6	1	272-408
999	RNase_PH_C	3' exoribonuclease family, domain 2	1.4e-13	58.5	2	96-157:392-469
999	PNPase	Polyribonucleotide nucleotidyltransferase	8.1e-05	25.9	1	189-269
1000	ig	Immunoglobulin domain	0.00054	26.7	1	38-102
1002	zf-C2H2	Zinc finger, C2H2 type	7.6e-54	192.3	7	183-205:211-233:239-261:267-289:295-317:323-345:351-373
1002	zf-BED	BED zinc finger	3.9	-3.7	1	224-262

Table 4B  
472

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1002	LIM	LIM domain	7.8	-19.2	1	185-249
1003	zf-C2H2	Zinc finger, C2H2 type	2.3e-114	393.3	16	260-282:297-319:325-347:353-375:381-403:409-431:437-459:465-487:493-515:521-543:549-571:577-599:605-627:633-655:661-683:689-711
1003	zf-BED	BED zinc finger	0.41	5.0	1	366-404
1003	TFIIS	Transcription factor S-II (TFIIS)	0.75	2.4	2	353-391:468-503
1003	GATA	GATA zinc finger	1.8	-6.6	1	435-481
1003	PHD	PHD-finger	5.5	-16.7	1	326-389
1003	E6	Early Protein (E6)	6.9	-67.7	1	467-554
1003	Bowman-Birk leg	Bowman-Birk serine protease inhibitor	8	-18.8	1	355-411
1007	LIM	LIM domain	1.3e-38	141.7	3	390-448:450-507:510-576
1007	E7	E7 protein, Early protein	0.7	-39.5	1	235-422
1008	Transposase_8	Transposase	2.9	-12.7	1	53-148
1010	zf-C2H2	Zinc finger, C2H2 type	4.8e-06	33.5	3	2-24:29-52:154-177
1010	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	7.1	-7.0	1	4-34
1011	60s_ribosomal	60s Acidic ribosomal protein	1.2e-20	82.1	1	1-89
1012	abhydrolase_2	Phospholipase/Carboxylesterase	1.3e-24	95.2	1	9-213
1016	Anti_proliferat	BTG1 family	1.7e-73	257.6	1	11-253
1017	zf-C2H2	Zinc finger, C2H2 type	9.8e-71	248.4	11	147-169:213-231:269-291:325-347:353-375:381-403:409-431:437-459:465-487:493-515:521-543
1017	KRAB	KRAB box	2.6e-26	100.8	1	42-82
1017	DM-domain	DM DNA binding domain	0.26	0.8	1	491-539
1017	zf-BED	BED zinc finger	1.8	-0.6	2	338-376:394-432
1017	LIM	LIM domain	3.4	-16.2	1	495-549
1017	zf-TRAF	TRAF-type zinc finger	5	-6.6	1	427-480
1017	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	9.5	-8.1	1	177-218
1017	Bowman-Birk leg	Bowman-Birk serine protease inhibitor	9.6	-19.6	1	147-203
1020	LIM	LIM domain	2.3e-34	127.6	4	78-136:139-197:200-256:259-315
1020	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	3.3	-4.2	1	106-142
1020	TNFR_c6	TNFR/NGFR cysteine-rich region	8.2	-0.1	1	89-127



Table 4B  
473

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1024	zf-C2H2	Zinc finger, C2H2 type	4.3e-53	189.8	10	156-183:184-206:212-234:240-262:268-290:296-318:324-346:352-374:380-402:408-430
1024	KRAB	KRAB box	2.7e-26	100.8	1	15-55
1024	TFIIS	Transcription factor S-II (TFIIS)	2.5	-2.0	1	129-166
1024	zf-BED	BED zinc finger	3.9	-3.8	1	225-263
1024	fer4	4Fe-4S binding domain	8.3	-3.2	1	176-199
1025	RAG2	Recombination activating protein 2	0	1380.2	1	1-527
1025	PHD	PHD-finger	2.9	-14.2	1	419-483
1026	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1.1e-06	35.6	1	16-54
1029	Keratin_B2	Keratin, high sulfur B2 protein	3.8e-06	16.4	2	5-149:150-281
1031	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-tr	4.1e-107	369.3	1	5-165
1035	zf-C2H2	Zinc finger, C2H2 type	1.2e-113	391.0	17	100-122:132-154:160-182:188-210:216-238:244-266:272-294:300-322:328-350:356-378:384-406:412-434:440-462:468-490:496-518:524-546:552-574
1035	KRAB	KRAB box	7.1e-24	92.8	1	4-44
1035	zf-BED	BED zinc finger	0.25	6.9	2	78-123:537-575
1035	LIM	LIM domain	1.5	-13.2	1	330-394
1035	TFIIS	Transcription factor S-II (TFIIS)	2.4	-1.9	1	244-282
1035	zf-MIZ	MIZ zinc finger	8.5	-20.9	1	457-507
1038	Sulfate_transp	Sulfate transporter family	4.3e-103	355.9	2	2-284:441-751
1038	STAS	STAS domain	4.8e-20	80.0	1	774-987
1038	BenE	Benzoate membrane transport protein	1.3	-317.8	1	422-745
1038	UPF0118	Domain of unknown function DUF20	4.9	-125.1	1	454-754
1038	Transp_cyt_pur	Permease for cytosine/purines, uracil	5.3	-195.7	1	403-760
1038	voltage_CLC	Voltage gated chloride channel	5.5	-238.5	1	378-680
1038	FecCD	FecCD transport family	7.9	-219.3	1	440-686
1039	ubiquitin	Ubiquitin family	1.9e-08	41.5	1	55-126
1042	ig	Immunoglobulin domain	3.1e-30	113.9	5	62-129:163-229:264-316:349-400:433-501
1043	UPAR_LY6	u-PAR/Ly-6 domain	7.7e-31	115.9	1	88-158
1044	7tm_5	7TM chemoreceptor	0.17	-139.3	1	29-293
1044	sugar_tr	Sugar (and other) transporter	0.27	-153.3	1	1-419

Table 4B  
474

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1044	Competence	Competence protein	1.8	-92.8	1	10-207
1044	oxidored_q1	NADH-Ubiquinone/plastoquinone (complex I)	2.6	-160.0	1	185-410
1044	PNTB	NAD(P) transhydrogenase beta subunit	6.6	-392.2	1	188-508
1044	COX1	Cytochrome C and Quinol oxidase polypepti	7.4	-311.2	1	1-396
1045	Lipase_chap	Proteobacterial lipase chaperone prot	3.1	-124.4	1	95-336
1045	Cys_Met_Me ta_PP	Cys/Met metabolism PLP-dependent enzy	8.2	-317.8	1	40-279
1046	PAP2	PAP2 superfamily	7.7	-31.4	1	88-175
1047	MHC_I	Class I Histocompatibility antigen, domains	7.9e-148	504.5	1	25-203
1047	ig	Immunoglobulin domain	0.79	16.0	1	220-285
1050	ig	Immunoglobulin domain	8.7e-07	36.0	1	38-110
1051	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.0019	-4.7	1	4-194
1052	GASA	Gibberellin regulated protein	2.6	-47.4	1	1-69
1054	SRCR	Scavenger receptor cysteine-rich domain	6.2e-25	96.3	1	722-820
1055	P2X_receptor	ATP P2X receptor	8.7e-303	1019.3	1	13-388
1055	Metallothio_5	Metallothionein family 5	7.1	-0.8	1	124-161
1057	zona_pellucida	Zona pellucida-like domain	3.1e-81	283.3	1	268-538
1057	trefoil	Trefoil (P-type) domain	0.02	9.1	1	224-262
1058	Aa_trans	Transmembrane amino acid transporter prot	2.4e-22	87.7	1	4-388
1058	oxidored_q1	NADH-Ubiquinone/plastoquinone (complex I)	2.6	-159.9	1	120-362
1058	DUF286	Caenorhabditis protein of unknown functio	2.7	-94.1	1	83-246
1058	PUCC	PUCC protein	8.4	-281.4	1	83-365
1058	Competence	Competence protein	8.4	-105.8	1	99-341
1059	7tm_1	7 transmembrane receptor (rhodopsin family)	1.5e-05	10.6	1	33-401
1060	Na_sulph_sym	Sodium:sulfate symporter transmembran	4.5e-143	488.7	1	1-554
1060	Na_H_antipor	Na+/H+ antiporter family	5.8	-114.4	1	212-546
1060	LrgA	LrgA family	7.6	-55.7	1	37-141
1061	oxidored_q1	NADH-Ubiquinone/plastoquinone (complex I)	9.5	-171.4	1	129-355
1063	MHC_I	Class I Histocompatibility antigen, domains	9.3e-06	-8.4	1	29-204
1064	ERG4_ERG24	Ergosterol biosynthesis ERG4/ERG24 family	1.1e-103	357.8	1	7-350
1065	Sulfatase	Sulfatase	0.28	-144.2	1	66-431
1065	CRCB	CrcB-like protein	2.6	-51.2	1	728-827
1065	Phosphodiester	Type I phosphodiesterase /	2.7	-167.6	1	43-368

Table 4B  
475

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		nucleotide p				
1065	Cad	Cadmium resistance transporter	7.5	-136.9	1	412-565
1065	Metalloenzyme	Metalloenzyme superfamily	9.3	-43.8	1	189-308
1066	ig	Immunoglobulin domain	1.2e-15	65.4	3	42-95:135-192:231-288
1067	7tm_1	7 transmembrane receptor (rhodopsin family)	1.5e-28	108.3	1	28-266
1068	UPF0136	Uncharacterised protein family (UPF0136)	4e-45	163.4	1	12-108
1069	pkinase	Protein kinase domain	1.6e-65	231.1	1	298-580
1069	Activin_recp	Activin types I and II receptor domain	6.2e-28	106.2	1	20-107
1069	toxin	Snake toxin	2.8	-19.3	1	33-101
1069	UPAR_LY6	u-PAR/Ly-6 domain	4	-19.7	1	47-101
1070	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.00018	13.2	1	3-177
1070	UPF0136	Uncharacterised protein family	3.6	-32.2	1	67-164
1071	MHC_I	Class I Histocompatibility antigen, domains	1.2e-12	55.5	1	24-202
1071	ig	Immunoglobulin domain	3.6e-05	30.6	1	218-284
1072	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	3.1	-60.4	1	36-207
1073	sugar_tr	Sugar (and other) transporter	0.032	-126.5	1	48-528
1073	PUCC	PUCC protein	4.5	-274.8	1	130-432
1073	Competence	Competence protein	6.5	-103.7	1	115-316
1073	Nuc_H_symp ort	Nucleoside H <sup>+</sup> symporter	9.6	-276.3	1	72-523
1074	sugar_tr	Sugar (and other) transporter	0.032	-126.5	1	136-616
1074	PUCC	PUCC protein	4.5	-274.8	1	218-520
1074	Competence	Competence protein	6.5	-103.7	1	203-404
1074	Nuc_H_symp ort	Nucleoside H <sup>+</sup> symporter	9.6	-276.3	1	160-611
1075	ank	Ankyrin repeat	1.7e-45	164.6	6	31-63:64-96:97-129:130-162:163-195:196-228
1076	7tm_1	7 transmembrane receptor (rhodopsin family)	5.7e-05	-2.6	1	118-374
1076	DUF216	Domain of unknown function DUF	1.6	-42.6	1	193-368
1077	sugar_tr	Sugar (and other) transporter	0.0039	-100.2	1	46-470
1077	OATP_C	Organic Anion Transporter Polypeptide (OATP)	0.018	-180.1	1	64-361
1077	PUCC	PUCC protein	6.2	-278.1	1	85-480
1077	PAP2	PAP2 superfamily	7.9	-31.5	1	353-476
1078	UvrD-helicase	UvrD/REP helicase	7.1	-248.3	1	299-774
1078	DUF246	Plant protein family	9.1	-257.7	1	584-865

Table 4B  
476

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1079	G_glu_transp ept	Gamma-glutamyltranspeptidase	3.2e-05	-146.3	1	122-499
1079	DLH	Dienelactone hydrolase family	2.5	-92.8	1	189-376
1079	DUF312	Short repeats of unknown function	3.3	0.8	1	415-453
1080	TPR	TPR Domain	2e-19	78.0	5	28-61:68-101:108-141:148-181:188-221
1081	ACAT	Sterol O-acyltransferase	1.9e-32	121.3	1	300-406
1082	WD40	WD domain, G-beta repeat	2.5e-20	81.0	7	1015-1050:1059-1097:1115-1151:1158-1194:1203-1240:1246-1281:1293-1329
1082	V1R	Vomeranosal organ pheromone receptor family	3.2	-162.0	1	421-676
1083	pkinase	Protein kinase domain	1.4e-54	194.8	1	248-568
1083	Activin_recp	Activin types I and II receptor domain	5.8e-38	139.6	1	26-127
1083	UPAR_LY6	u-PAR/Ly-6 domain	7.9	-22.8	1	48-121
1084	pkinase	Protein kinase domain	1.4e-54	194.8	1	310-630
1084	Activin_recp	Activin types I and II receptor domain	5.8e-38	139.6	1	26-127
1084	UPAR_LY6	u-PAR/Ly-6 domain	7.9	-22.8	1	48-121
1086	C2	C2 domain	1.7e-06	35.0	1	233-316
1087	Cache	Cache domain	2.7e-24	94.2	1	557-650
1087	vwa	von Willebrand factor type A domain	0.13	-30.6	1	361-537
1088	Glyco_hydro 31	Glycosyl hydrolases family 31	1.6e-224	759.3	1	1-636
1090	SSrecog	Structure-specific recognition protein	2.7e-153	522.7	1	343-555
1090	HMG_box	HMG (high mobility group) box	3.6e-32	120.3	1	681-749
1090	DUF304	Bacterial membrane flanked domain	4.8	-12.0	1	144-223
1091	serpin	Serpin (serine protease inhibitor)	8.1e-196	663.9	1	315-683
1092	trypsin	Trypsin	0.00023	-4.5	1	334-526
1093	Competence	Competence protein	9.1	-106.5	1	9-260
1094	Aa_trans	Transmembrane amino acid transporter pro	1.5e-07	4.2	1	115-561
1094	Trp_Tyr_per m	Tryptophan/tyrosine permease family	2.7	-283.0	1	118-558
1094	oxidored_q1	NADH-Ubiquinone/plastoquinone	8.1	-170.0	1	117-338
1097	ATP-synt_C	ATP synthase subunit C	3.3e-18	73.9	1	72-140
1098	aconitase	Aconitase family (aconitate hydratase)	1.9e-177	602.9	1	162-744
1098	Aconitase_C	Aconitase C-terminal domain	8.9e-72	251.9	1	872-1043
1103	PAP2	PAP2 superfamily	6.3e-15	63.0	1	89-236
1104	ig	Immunoglobulin domain	1.3e-18	75.3	2	80-148:183-251

Table 4B  
477

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1106	TCTP	Translationally controlled tumor protein	3.5e-93	323.0	1	1-166
1109	efhand	EF hand	1.3e-13	58.7	3	33-61:102-130:138-166
1110	ATP1G1_PL M_MAT8	ATP1G1/PLM/MAT8 family	1.8e-13	58.2	1	92-146
1112	ldl_recept_a	Low-density lipoprotein receptor domain	0.00073	26.3	1	115-153
1112	CUB	CUB domain	0.0032	-3.1	1	9-109
1115	7tm_1	7 transmembrane receptor (rhodopsin family)	1.8e-05	9.0	1	139-349
1116	7tm_1	7 transmembrane receptor	7.5e-21	82.7	1	114-363
1116	oxidored_q5_N	NADH-ubiquinone oxidoreductase chain 4	4.2	-55.0	1	213-282
1116	Abi	CAAX amino terminal protease family	9.4	-28.2	1	171-254
1116	PhaG_MnhG_YufB	Na <sup>+</sup> /H <sup>+</sup> antiporter subunit	9.8	-57.7	1	281-376
1119	tsp_1	Thrombospondin type 1 domain	4.5e-38	139.9	10	149-198:306-364:571-626:631-696:707-761:841-889:970-1021:1099-1148:1219-1269:1342-1398
1120	HMG14_17	HMG14 and HMG17	1.3e-34	128.4	1	2-86
1124	lectin_c	Lectin C-type domain	5.1	-34.2	1	216-281
1125	ig	Immunoglobulin domain	2.1e-17	71.3	3	84-153:185-255:292-347
1136	sugar_tr	Sugar (and other) transporter	0.029	-125.4	1	23-504
1136	OATP_C	Organic Anion Transporter Polypeptide (OATP)	2.8	-225.4	1	116-381
1136	Pucc	Pucc protein	5.8	-277.4	1	83-476
1136	Frizzled	Frizzled/Smoothed family membrane region	8.3	-247.2	1	54-264
1137	ATP-synt_C	ATP synthase subunit C	3.3e-35	130.4	2	14-79:90-155
1138	DHDPS	Dihydrodipicolinate synthetase family	7.2e-31	116.0	1	34-325
1139	SSF	Sodium:solute symporter family	1.5e-48	174.7	2	50-461:569-953
1139	PNTB	NAD(P) transhydrogenase beta subunit	7.6	-393.3	1	62-347
1139	Competence	Competence protein	8.9	-106.3	1	64-287
1140	Cdc37	Cdc37 family	5.4	-122.6	1	317-539
1143	7tm_1	7 transmembrane receptor (rhodopsin family)	1.3e-44	161.6	1	78-332
1144	Peptidase_A8	Signal peptidase (SPase) II	7.2	-70.3	1	117-211
1145	PID	Phosphotyrosine interaction domain (PTB/PID)	1.7e-99	343.9	2	488-627:661-782
1145	WW	WW domain	1.7e-08	41.7	1	373-401
1145	KdpC	K <sup>+</sup> -transporting ATPase, c chain	3.6	-97.4	1	60-245
1146	PRA1	Prenylated rab acceptor (PRA1)	0.0071	-32.6	1	1-153

Table 4B  
478

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1147	DUF214	Predicted permease	9.5	-68.5	1	17-151
1148	C2	C2 domain	2.1e-30	114.4	2	87-165:240-320
1149	Reticulon	Reticulon	0.29	-68.6	1	65-266
1150	Reticulon	Reticulon	1.8	-81.1	1	79-266
1152	7tm_1	7 transmembrane receptor	2.8e-27	104.1	1	161-410
1152	oxidored_q1	NADH-Ubiquinone/plastoquinone (complex I)	5.3	-166.1	1	90-329
1153	WD40	WD domain, G-beta repeat	9.8e-06	32.5	4	26-62:73-109:236-271:430-467
1155	aa_permeases	Amino acid permease	5.6e-06	-179.9	1	193-613
1155	7tm_5	7TM chemoreceptor	1.6	-156.9	1	199-448
1155	oxidored_q1	NADH-Ubiquinone/plastoquinone	3.7	-162.9	1	306-552
1155	Spore_permease	Spore germination protein	5.6	-193.9	1	180-533
1155	Aa_trans	Transmembrane amino acid transporter p	6.4	-205.1	1	179-589
1155	BPD_transp_2	Branched-chain amino acid transport sy	6.5	-123.5	1	47-327
1157	PRA1	Prenylated rab acceptor (PRA1)	1.1e-46	168.5	1	23-149
1158	dUTPase	dUTPase	5.1e-29	109.8	1	46-167
1158	rvp	Retroviral aspartyl protease	4.1e-22	86.9	1	179-280
1158	G-patch	G-patch domain	0.0011	25.7	1	285-329
1158	MIP	Major intrinsic protein	0.048	-83.0	1	1-111
1159	synaptobrevin	Synaptobrevin	0.49	-3.1	1	96-184
1160	laminin_EGF	Laminin EGF-like (Domains III and V)	9.5e-132	451.1	14	299-353:356-423:426-466:490-532:535-578:584-627:630-680:683-727:1265-1308:1311-1352:1355-1401:1404-1452:1686-1730:1733-1786
1160	laminin_Nterm	Laminin N-terminal (Domain VI)	4e-77	269.6	1	47-297
1160	laminin_G	Laminin G domain	6e-31	116.2	3	2625-2748:3015-3138:3185-3313
1160	laminin_B	Laminin B (Domain IV)	1.1e-09	45.6	1	1517-1651
1160	MCPsignal	Methyl-accepting chemotaxis protein	0.61	-119.3	1	2011-2237
1160	TSPN	Thrombospondin N-terminal -like domain	1.7	-47.9	1	3143-3311
1160	filament	Intermediate filament protein	2.2	-193.4	1	1826-2081
1160	DSL	Delta serrate ligand	3.3	-17.5	1	472-519
1160	OEP	Outer membrane efflux protein	3.7	-29.3	1	2020-2211
1160	EB	EB module	6.5	-10.6	1	665-713
1160	Apidaecin	Apidaecin	7	7.3	1	907-934
1160	Acetate_kinase	Acetokinase family	8.3	-242.2	1	2093-2412

Table 4B  
479

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1160	spectrin	Spectrin repeat	8.4	-20.0	1	2105-2205
1161	EGF	EGF-like domain	2.2e-26	101.1	5	29-57:60-88:104-140:147-178:185-216
1161	EB	EB module	1.7	-4.2	1	131-178
1161	metalthio	Metallothionein	7.7	-11.7	1	64-133
1162	interferon	Interferon alpha/beta domain	1.5e-17	71.7	1	16-171
1163	IL1	Interleukin-1 / 18	8.7e-21	82.5	1	19-159
1164	PDGF	Platelet-derived growth factor (PDGF)	1.8e-51	184.4	1	52-130
1165	IL1	Interleukin-1 / 18	3.7e-23	90.4	1	11-144
1167	Palm_thioest	Palmitoyl protein thioesterase	2.4e-216	732.1	1	28-282
1168	serpin	Serpin (serine protease inhibitor)	1.6e-202	686.2	1	47-415
1170	Defensin_pro pep	Defensin propeptide	1.2e-26	102.0	1	38-90
1170	defensins	Mammalian defensin	7.3e-14	59.5	1	103-131
1172	Y_phosphatase	Protein-tyrosine phosphatase	1.8e-110	380.4	1	271-499
1179	Metallophos	Calcineurin-like phosphoesterase	1.1e-11	52.3	1	70-285
1184	MHC_I	Class I Histocompatibility antigen, domains	6.2e-06	-5.9	1	29-205
1187	PLA2_B	Lysophospholipase catalytic domain	8.3e-53	188.9	1	357-800
1187	C2	C2 domain	1.6e-06	35.1	1	46-129
1189	lipocalin	Lipocalin / cytosolic fatty-acid binding	2.6e-39	144.0	1	39-188
1189	Kunitz_BPTI	Kunitz/Bovine pancreatic trypsin inhibitor	4.7e-23	90.0	2	231-273:274-319
1190	ig	Immunoglobulin domain	4e-07	37.1	1	38-116
1191	ig	Immunoglobulin domain	4e-07	37.1	1	38-116
1196	UPF0185	Uncharacterised protein family (UPF0185)	3.6e-41	150.2	1	15-102
1200	zf-DHHC	DHHC zinc finger domain	3.7e-15	63.8	1	87-149
1201	KRAB	KRAB box	6.6e-22	86.2	1	33-73
1202	KRAB	KRAB box	6.6e-22	86.2	1	33-73
1204	trypsin	Trypsin	3e-31	117.2	1	56-239
1208	lipocalin	Lipocalin / cytosolic fatty-acid binding pr	0.12	9.2	1	38-186
1212	LRR	Leucine Rich Repeat	2.5e-54	193.9	15	73-96:97-122:123-149:150-175:176-195:197-220:221-246:247-270:292-317:318-337:339-362:363-388:389-408:410-433:434-459
1212	LRRNT	Leucine rich repeat N-terminal domain	0.24	17.9	1	43-71
1212	SCAN	SCAN domain	7.2	-38.8	1	88-187
1213	Zn_carbOpept	Zinc carboxypeptidase	1.9e-61	217.6	1	50-285

Table 4B  
480

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1214	ig	Immunoglobulin domain	5.5	9.1	1	17-84
1216	death	Death domain	2.4	-0.1	1	93-180
1217	death	Death domain	1.5	2.0	1	73-160
1221	C1q	C1q domain	3.4e-41	150.3	1	134-258
1221	Collagen	Collagen triple helix repeat (20 copies)	2.6e-10	47.7	1	41-100
1223	pkinase	Protein kinase domain	9e-75	261.8	1	158-417
1223	SH3	SH3 domain	0.019	7.4	1	423-478
1223	UBA	UBA/TS-N domain	0.035	20.7	1	987-1026
1224	vwa	von Willebrand factor type A domain	1.2e-05	29.0	1	51-242
1225	abhydrolase	alpha/beta hydrolase fold	3.7e-12	53.8	1	111-390
1225	abhydrolase_2	Phospholipase/Carboxylesterase	2.3	-88.8	1	64-270
1227	lectin_c	Lectin C-type domain	9.6e-29	108.9	1	54-166
1237	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	1.5e-09	-1.3	1	242-385
1257	LRR	Leucine Rich Repeat	9e-65	228.6	14	78-101:102-125:126-149:150-173:174-197:198-221:222-245:246-269:270-293:294-317:318-341:342-365:366-389:390-413
1257	LRRCT	Leucine rich repeat C-terminal domain	0.011	22.3	1	423-475
1257	LRRNT	Leucine rich repeat N-terminal domain	0.23	18.0	1	24-52
1263	Pep_M12B_p ropep	Reprolysin family propeptide	1.8e-28	108.1	1	75-191
1263	disintegrin	Disintegrin	2.7e-11	51.0	1	419-494
1263	Reprolysin	Reprolysin (M12B) family zinc metallo	7.3e-07	-22.2	1	206-402
1263	EGF	EGF-like domain	2.7	10.3	1	635-663
1266	Clusterin	Clusterin	2.3e-298	1004.6	1	2-394
1267	serpin	Serpin (serine protease inhibitor)	3.8e-140	479.0	1	80-425
1268	serpin	Serpin (serine protease inhibitor)	6.8e-162	551.2	1	98-461
1269	COX6C	Cytochrome c oxidase subunit VIc	2.3e-38	140.9	1	1-75
1270	Cadherin_C_term	Cadherin cytoplasmic region	4	-79.4	1	13-130
1271	spectrin	Spectrin repeat	8.2	-19.9	1	307-417
1272	W2	eIF4-gamma/eIF5/eIF2-epsilon	7.3	-26.9	1	325-395
1272	spectrin	Spectrin repeat	8.3	-19.9	1	327-437
1273	Rhodanese	Rhodanese-like domain	1.5e-48	174.7	2	16-138:165-266
1274	trypsin	Trypsin	2.9e-80	280.1	1	121-346
1279	Calpain_III	Calpain large subunit, domain III	1.5e-25	98.3	1	30-179
1284	Ribosomal_L_35p	Ribosomal protein L35	2.4	-18.1	1	122-179
1285	Ribosomal_L	Ribosomal protein L35	2.4	-18.1	1	122-179



Table 4B  
481

SEQ ID	Model	Description	E-value	Score	Repeats	Position
	35p					
1287	zf-C2H2	Zinc finger, C2H2 type	4.4e-66	233.0	9	173-195:201-223:229-251:257-279:285-307:313-335:341-363:369-391:397-419
1287	KRAB	KRAB box	4.9e-23	90.0	1	14-54
1287	zf-BED	BED zinc finger	2.4	-1.9	1	186-224
1287	LIM	LIM domain	9.3	-19.9	1	175-239
1292	EGF	EGF-like domain	0.36	17.3	1	30-65
1294	vwc	von Willebrand factor type C domain	1.3e-09	45.4	1	69-127
1295	zf-C2H2	Zinc finger, C2H2 type	5e-50	179.6	8	263-285:319-341:347-369:375-397:403-425:431-453:459-481:487-510
1295	KRAB	KRAB box	8.9e-26	99.1	1	8-48
1295	TFIIS	Transcription factor S-II (TFIIS)	5.7	-5.0	1	431-469
1295	zf-TRAF	TRAF-type zinc finger	7.3	-8.4	1	369-419
1295	LIM	LIM domain	7.7	-19.2	1	321-385
1296	zf-C2H2	Zinc finger, C2H2 type	3.2e-113	389.6	13	190-212:218-240:246-268:274-296:302-324:330-352:358-380:386-408:414-436:442-464:470-492:498-520:526-548
1296	KRAB	KRAB box	6.7e-28	106.1	1	14-54
1296	zf-BED	BED zinc finger	0.51	4.1	1	511-549
1296	PHD	PHD-finger	2.4	-13.4	1	275-338
1296	TFIIS	Transcription factor S-II (TFIIS)	7.6	-6.1	1	330-368
1296	LIM	LIM domain	8.1	-19.4	1	388-456
1299	hemopexin	Hemopexin	7e-25	96.1	4	279-322:324-367:369-415:417-461
1299	Peptidase_M10	Matrixin	8.4e-14	58.8	1	1-176
1299	Astacin	Astacin (Peptidase family M12A)	1.4	-117.0	1	91-238
1301	ig	Immunoglobulin domain	0.67	16.4	1	33-116
1305	DUF25	Domain of unknown function DUF25	5.3e-76	265.9	1	40-162
1307	kazal	Kazal-type serine protease inhibitor domain	1.3e-12	55.3	1	21-68
1309	Defensin_pro pep	Defensin propeptide	1.2e-26	102.0	1	74-126
1309	defensins	Mammalian defensin	7.3e-14	59.5	1	139-167
1312	ras	Ras family	2.7e-59	210.4	1	64-253
1312	GTP_EFTU	Elongation factor Tu GTP binding domain	0.91	-79.4	1	61-226
1312	arf	ADP-ribosylation factor family	4.3	-116.7	1	46-226

Table 4B

482

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1314	zf-C2H2	Zinc finger, C2H2 type	1.1e-51	185.1	8	163-185:191-213:219-241:247-269:275-297:303-325:331-353:359-381
1314	LIM	LIM domain	1.8	-13.9	1	221-285
1314	PHD	PHD-finger	2.5	-13.5	1	304-367
1314	zf-BED	BED zinc finger	2.8	-2.4	1	232-270
1315	Keratin_B2	Keratin, high sulfur B2 protein	0.0035	-30.0	1	13-154
1316	Keratin_B2	Keratin, high sulfur B2 protein	3.6e-15	63.8	1	23-154
1317	Keratin_B2	Keratin, high sulfur B2 protein	8.2e-20	79.2	2	20-163:164-295
1318	Keratin_B2	Keratin, high sulfur B2 protein	5.5	-79.9	1	3-172
1319	zf-C2H2	Zinc finger, C2H2 type	2.9e-52	187.1	7	33-55:61-83:89-111:117-139:145-167:173-195:201-223
1319	PHD	PHD-finger	1.6	-11.8	1	146-209
1319	zf-BED	BED zinc finger	3.4	-3.2	1	74-112
1321	Keratin_B2	Keratin, high sulfur B2 protein	0.00014	-8.3	1	23-193
1331	ig	Immunoglobulin domain	9.7e-06	32.5	1	24-78
1332	Kunitz_BPTI	Kunitz/Bovine pancreatic trypsin inhibito	2.2e-18	74.5	1	76-126
1333	zf-C2H2	Zinc finger, C2H2 type	4.6e-55	196.3	7	199-221:227-249:255-277:283-305:311-333:339-361:367-389
1333	KRAB	KRAB box	5.8e-24	93.0	1	6-46
1333	zf-BED	BED zinc finger	0.31	6.1	2	240-278:352-390
1336	vwa	von Willebrand factor type A domain	0.0038	-8.1	1	32-193
1336	TIG	IPT/TIG domain	1.4	4.5	1	199-291
1337	vwa	von Willebrand factor type A domain	0.0038	-8.1	1	32-193
1337	TIG	IPT/TIG domain	1.4	4.5	1	199-291
1339	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.0088	-16.3	1	4-203
1340	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	0.0019	-4.7	1	4-194
1341	PsaL	Photosystem I reaction centre subunit XI	4.6	-111.8	1	21-106
1343	ig	Immunoglobulin domain	1.5e-13	58.5	2	81-158:204-278
1344	MARCKS	MARCKS family	6.7	-95.4	1	124-395
1347	ig	Immunoglobulin domain	1.2e-07	38.9	2	168-226:276-330
1348	ig	Immunoglobulin domain	1.5e-18	75.1	3	157-217:269-325:373-427
1348	rhv	picornavirus capsid protein	1.4	-52.0	1	28-258
1353	LRR	Leucine Rich Repeat	2e-39	144.4	9	120-143:144-167:168-191:192-

Table 4B  
483

SEQ ID	Model	Description	E-value	Score	Repeats	Position
						215:216-239:240-263:264-287:311-334:335-358
1353	LRRNT	Leucine rich repeat N-terminal domain	0.0066	23.1	1	92-119
1353	LRRCT	Leucine rich repeat C-terminal domain	1.5	4.9	1	368-418
1353	Bgal_small_C	Beta galactosidase small chain, C termin	6.6	-44.4	1	105-174
1354	oxidored_q4	NADH-ubiquinone/plastoquinone oxidoreduct	5.8	-35.2	1	60-142
1356	ig	Immunoglobulin domain	2.1e-17	71.3	3	37-106:138-208:245-300
1357	CD36	CD36 family	3.9e-231	781.3	1	5-445
1358	DUF139	Cysteine rich repeat (DUF139)	1.2e-08	42.2	4	195-211:245-261:288-304:316-332
1359	ig	Immunoglobulin domain	1.6e-06	35.1	2	41-124:156-230
1359	Gag_MA	Matrix protein (MA), p15	0.0044	-25.2	1	291-429
1360	ion_trans	Ion transport protein	0.0031	24.2	1	285-445
1360	DUF214	Predicted permease	8.6	-67.8	1	202-378
1363	spectrin	Spectrin repeat	9.5e-22	85.7	6	17-121:124-226:229-340:372-476:678-785:788-896
1363	Remorin_C	Remorin, C-terminal region	3.5	-39.8	1	659-744
1366	C1q	C1q domain	2.1e-31	117.8	1	266-390
1366	Collagen	Collagen triple helix repeat (20 copies)	0.00029	19.8	1	182-241
1367	rnaseA	Pancreatic ribonuclease	1.1e-34	128.7	1	27-132
1368	3HCDH_N	3-hydroxyacyl-CoA dehydrogenase, NAD binding	0.0021	-60.2	1	50-141
1369	3HCDH_N	3-hydroxyacyl-CoA dehydrogenase, NAD binding	0.25	-91.5	1	50-153
1371	voltage_CLC	Voltage gated chloride channel	3.1e-199	675.3	1	92-528
1371	CBS	CBS domain	4.9e-16	66.7	2	559-617:761-814
1371	PHO4	Phosphate transporter family	4.1	-199.1	1	129-495
1371	K_trans	K+ potassium transporter	5.8	-566.6	1	267-759
1371	7tm_5	7TM chemoreceptor	6.3	-167.6	1	254-486
1372	7tm_1	7 transmembrane receptor (rhodopsin family)	0.2	-84.7	1	82-199
1373	ldh_C	lactate/malate dehydrogenase, alpha/beta C-t	0.69	-61.5	1	38-200
1374	disintegrin	Disintegrin	7.2	-30.4	1	241-306
1375	Peptidase_C1	Papain family cysteine protease	1e-120	414.5	1	114-332
1376	7tm_1	7 transmembrane receptor	2.6e-50	180.6	1	48-454
1376	7tm_5	7TM chemoreceptor	1	-153.3	1	26-453
1376	Sec62	Translocation protein Sec62	1.4	-124.2	1	307-477

Table 4B  
484

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1376	DUF280	Caenorhabditis protein of unknown function	5.6	-141.0	1	93-274
1376	Sre	C. elegans Sre G protein-coupled chemoreceptor	7.2	-136.9	1	1-503
1376	Colicin_im	Colicin immunity protein	9.6	-61.3	1	365-466
1378	7tm_2	7 transmembrane receptor (Secretin family)	4.9e-09	32.5	1	250-500
1378	GPS	Latrophilin/CL-1-like GPS domain	0.38	4.4	1	197-248
1378	Competence	Competence protein	5.9	-102.8	1	240-490
1380	MAGE	MAGE family	1.4e-91	317.7	1	3-231
1381	ion_trans	Ion transport protein	0.065	16.8	1	158-323
1381	ABC2_membrane	ABC-2 type transporter	6.1	-130.1	1	150-339
1381	oxidored_q1	NADH-Ubiquinone/plastoquinone	6.2	-167.5	1	51-322
1381	OATP_C	Organic Anion Transporter Polypeptide	9.5	-236.4	1	80-321
1382	ig	Immunoglobulin domain	7e-12	52.9	2	37-128:160-241
1385	MHC_I	Class I Histocompatibility antigen, domains	9.3e-06	-8.4	1	29-204
1388	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-tr	7e-88	305.4	1	45-184
1390	7tm_1	7 transmembrane receptor (rhodopsin family)	5.3e-06	21.2	1	1-210
1391	Reticulon	Reticulon	6.4	-89.9	1	52-257
1392	Collagen	Collagen triple helix repeat	0.92	-48.2	1	27-84
1392	MDM2	p53-associated protein (MDM2)	5.1	-12.8	1	92-155
1392	Cytidyltransferase	Phosphatidate cytidyltransferase	9.4	-91.4	1	141-242
1393	7tm_1	7 transmembrane receptor (rhodopsin family)	1.6e-12	55.0	1	99-329
1393	DUF40	Domain of unknown function DUF40	5.7	-131.9	1	97-267
1395	trefoil	Trefoil (P-type) domain	3.9	-11.1	1	28-53
1397	fn3	Fibronectin type III domain	4.7e-78	272.7	5	533-619:631-717:734-833:854-939:951-1039
1397	ig	Immunoglobulin domain	3.7e-42	153.5	5	154-227:261-318:362-418:453-511:1253-1316
1398	DUF300	Domain of unknown function	1.4e-152	520.3	1	51-341
1404	ig	Immunoglobulin domain	2.9e-08	40.9	2	42-99:139-198
1409	ig	Immunoglobulin domain	8.1e-30	112.5	3	142-199:263-321:359-438
1410	Tissue_fac	Tissue factor	0.26	-135.2	1	1-267
1410	HRM	Hormone receptor domain	8.4	-13.9	1	28-101
1412	FecCD	FecCD transport family	7.2	-218.5	1	144-342
1413	Metallothio_Pro	Prokaryotic metallothionein	6.2	-8.0	1	464-511
1414	mito_carr	Mitochondrial carrier protein	3.5e-64	226.7	3	40-130:137-227:238-322
1416	ig	Immunoglobulin domain	0.0011	25.6	1	57-112

Table 4B  
485

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1417	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	6.6e-06	33.1	1	27-67
1418	Synaptophysin	Synaptophysin / synaptoporin	6.9e-71	248.9	1	27-272
1419	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	5.2e-18	73.3	1	4-165
1421	Ammonium_transporter	Ammonium Transporter Family	8.5e-18	72.6	1	25-383
1421	FecCD	FecCD transport family	1.8	-206.7	1	97-311
1421	Bac_chlorC	Bacteriochlorophyll C binding protein	5.8	-12.0	1	212-289
1423	rnaseA	Pancreatic ribonuclease	4.1e-29	110.1	1	32-136
1424	PCI	PCI domain	8.5	-9.1	1	64-133
1424	Nrap	Nrap protein	9.7	-257.5	1	162-764
1425	CbiM	CbiM	3.4	-86.8	1	16-138
1427	LRR	Leucine Rich Repeat	4.5e-26	100.0	6	3-26:27-50:51-74:76-99:100-123:124-144
1427	fn3	Fibronectin type III domain	2.7e-08	41.0	1	334-417
1427	ig	Immunoglobulin domain	2.4e-07	37.9	1	225-283
1427	LRRCT	Leucine rich repeat C-terminal domain	7.2e-06	32.9	1	164-209
1428	GRAM	GRAM domain	0.067	13.1	1	177-245
1429	YGGT	YGGT family	7.4	-24.4	1	147-217
1431	Papilloma_E5	Papillomavirus E5	6.7	-34.9	1	52-108
1435	Rhomboid	Rhomboid family	1.5e-45	164.8	1	169-326
1435	efhand	EF hand	3.8	11.0	1	38-66
1437	Sema	Sema domain	7e-202	684.1	1	102-513
1437	tsp_1	Thrombospondin type 1 domain	1.1e-42	155.3	6	589-637:644-695:702-746:833-883:890-940:945-985
1437	PSI	Plexin repeat	2.5e-14	61.0	1	531-578
1438	ank	Ankyrin repeat	2e-15	64.7	3	336-368:383-418:468-498
1438	ion_trans	Ion transport protein	9.4e-07	35.9	1	615-817
1439	ank	Ankyrin repeat	3.3e-12	54.0	3	417-449:464-499:549-579
1439	ion_trans	Ion transport protein	9.4e-07	35.9	1	696-898
1442	LRR	Leucine Rich Repeat	3.2e-20	80.6	5	57-80:81-104:105-128:129-152:153-176
1442	LRRCT	Leucine rich repeat C-terminal domain	2.3e-07	37.9	1	186-240
1442	LRRNT	Leucine rich repeat N-terminal domain	0.00013	28.8	1	27-55
1443	7tm_2	7 transmembrane receptor (Secretin family)	7.2e-23	89.4	1	166-476
1443	HMG_box	HMG (high mobility group) box	3.1e-05	30.8	1	3-76
1443	GPS	Latrophilin/CL-1-like GPS domain	0.0066	19.1	1	110-162
1443	7tm_1	7 transmembrane receptor (rhodopsin family)	1.9	-107.5	1	255-473

Table 4B  
486

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1443	Frizzled	Frizzled/Smoothed family membrane region	3.3	-241.3	1	156-459
1444	DUF214	Predicted permease	0.0051	-16.4	1	136-275
1444	secY	eubacterial secY protein	1.1	-230.8	1	13-270
1448	ribonuc_red_s m	Ribonucleotide reductase, small chain	5.6	-142.1	1	529-805
1452	mito_carr	Mitochondrial carrier protein	8.4e-58	205.5	2	27-123:124-217
1454	mito_carr	Mitochondrial carrier protein	0.00049	22.3	2	79-179:198-287
1455	WD40	WD domain, G-beta repeat	2.7e-26	100.8	4	64-99:105-141:235-273:279-315
1456	Furin-like	Furin-like cysteine rich region	6.1e-89	308.9	1	172-328
1456	Recep_L_do main	Receptor L domain	3.1e-80	280.0	2	46-169:345-471
1456	pkinase	Protein kinase domain	7e-65	229.0	1	986-1258
1456	fn3	Fibronectin type III domain	0.2	11.5	2	604-794:814-909
1456	fer4	4Fe-4S binding domain	3.7	-1.0	1	238-267
1456	Keratin_B2	Keratin, high sulfur B2 protein	5.2	-79.5	1	143-285
1456	TIL	Trypsin Inhibitor like cysteine rich d	6.3	-13.2	1	221-279
1462	Peptidase_M1 0	Matrixin	1.5e-79	277.7	1	62-227
1462	hemopexin	Hemopexin	1.4e-10	48.6	1	309-350
1463	Armadillo_se g	Armadillo/beta-catenin-like repeat	0.028	21.0	4	549-591:641-685:686-737:738-783
1466	C2	C2 domain	6.5e-07	36.4	1	219-304
1468	RTC	RNA 3'-terminal phosphate cyclase	6.1e-05	-118.8	1	9-196
1476	CN_hydrolas e	Carbon-nitrogen hydrolase	0.0014	-76.2	1	64-341
1477	zf-C2H2	Zinc finger, C2H2 type	4.8e-06	33.5	3	267-289:294-317:419-442
1477	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	7.1	-7.0	1	269-299
1478	Euk_porin	Eukaryotic porin	3.3e-06	-42.1	1	1-142
1482	ras	Ras family	0.24	-121.5	1	389-513
1484	BTB	BTB/POZ domain	2.3e-13	57.8	1	88-194
1487	rvt	Reverse transcriptase	8.1e-57	202.2	1	241-507
1489	Galactosyl_T 2	Galactosyltransferase	4.7e-177	601.6	1	97-404
1489	TT_ORF2	TT viral ORF2	9	-98.0	1	339-453
1490	Orexin	Prepro-orexin	8.3	-32.3	1	1-128
1491	GBP_C	Guanylate-binding protein, C-terminal domain	0.00059	-84.4	1	2-155
1493	NUDIX	NUDIX domain	0.0011	21.7	1	297-439
1496	Big_2	Bacterial Ig-like domain (group 2)	4.5e-16	66.8	1	1177-1258
1497	ig	Immunoglobulin domain	1.9e-10	48.2	1	38-112
1498	aldo_ket_red	Aldo/keto reductase family	3.7e-08	-8.7	1	108-367
1502	Caveolin	Caveolin	1.6e-06	13.3	1	90-262

Table 4B  
487

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1504	sushi	Sushi domain (SCR repeat)	8.2e-16	66.0	2	223-281:286-344
1511	zf-C2H2	Zinc finger, C2H2 type	1.1e-16	68.9	2	341-363:369-391
1511	KRAB	KRAB box	4.9e-09	43.5	1	230-268
1511	zf-BED	BED zinc finger	4.1	-3.9	1	354-392
1516	CBS	CBS domain	1.4e-30	115.0	3	189-243:264-317:336-389
1517	zf-C2H2	Zinc finger, C2H2 type	2.4e-128	439.8	18	279-301:307-329:335-357:363-385:391-413:419-441:447-469:475-497:503-525:868-890:896-918:924-946:952-974:980-1002:1008-1030:1036-1058:1064-1086:1092-1114
1517	KRAB	KRAB box	5.6e-40	146.2	2	74-114:681-713
1517	PHD	PHD-finger	1.2	-10.7	1	897-960
1517	SCAN	SCAN domain	1.6	-32.1	1	544-624
1517	LIM	LIM domain	5.4	-17.9	1	337-401
1517	BolA	BolA-like protein	6.1	-25.1	1	283-347
1517	TFIIS	Transcription factor S-II (TFIIS)	9.1	-6.8	1	952-990
1518	ig	Immunoglobulin domain	1e-06	35.7	1	38-102
1521	CoA_trans	Coenzyme A transferase	1.1e-161	550.5	2	42-261:271-472
1521	TGF-beta	Transforming growth factor beta like domain	6.8e-64	225.7	1	637-741
1523	Reprolysin	Reprolysin (M12B) family zinc metallo	6.6e-95	328.7	1	223-422
1523	Pep_M12B_p ropep	Reprolysin family propeptide	2e-27	104.5	1	121-248
1523	disintegrin	Disintegrin	1.9e-14	61.5	1	437-513
1523	EGF	EGF-like domain	0.0092	22.6	1	661-693
1523	VSP	Giardia variant-specific surface prot	4.9	-258.5	1	428-750
1523	Keratin_B2	Keratin, high sulfur B2 protein	7.2	-81.7	1	427-538
1524	DSPc	Dual specificity phosphatase, catalytic doma	3.6e-71	249.9	1	102-241
1527	Spermine_syn th	Spermine/spermidine synthase	8.2e-13	56.0	1	254-440
1528	efhand	EF hand	3.4e-10	47.3	4	65-91:145-173:329-357:365-393
1529	Epimerase	NAD dependent epimerase/dehydratase family	0.66	-192.0	1	10-451
1530	rvt	Reverse transcriptase	7.7e-69	242.1	1	125-395
1531	RPH3A_effec tor	Rabphilin-3A effector domain	1.9e-09	-6.1	1	161-339
1533	TPR	TPR Domain	3.7e-10	47.2	3	220-253:444-477:478-511
1533	CENP-B	CENP-B protein	1.2e-06	-7.7	1	541-671
1534	SAM	SAM domain (Sterile alpha	0.045	18.0	1	688-752

Table 4B  
488

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		motif)				
1534	Atrophin-1	Atrophin-1 family	7.9	-679.9	1	731-1402
1538	pkinase	Protein kinase domain	1.5e-20	81.7	1	42-242
1540	zf-C2H2	Zinc finger, C2H2 type	0.00035	27.3	1	725-750
1541	COLFI	Fibrillar collagen C-terminal domain	3.9e-163	555.3	1	624-841
1541	Collagen	Collagen triple helix repeat (20 copies)	2.2	-55.6	1	538-597
1541	TT_ORF2	TT viral ORF2	3.9	-93.3	1	508-610
1542	EGF	EGF-like domain	2.6e-11	51.0	3	916-947:1004-1044:1050-1094
1542	granulin	Granulin	6.3	-17.2	1	1019-1074
1542	TIL	Trypsin Inhibitor like cysteine rich domain	9.4	-15.2	1	938-1004
1543	ig	Immunoglobulin domain	1.5e-10	48.5	1	38-112
1544	UCH-1	Ubiquitin carboxyl-terminal hydrolases famil	8.5e-15	62.6	1	477-508
1544	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	7.3e-13	56.2	1	759-811
1544	zf-UBP	Zn-finger in ubiquitin-hydrolases and other	0.011	20.7	1	350-420
1545	spectrin	Spectrin repeat	2.8	-14.5	1	696-803
1546	LRR	Leucine Rich Repeat	1.1e-30	115.3	9	132-155:156-179:180-199:201-224:250-269:271-294:295-318:319-341:342-365
1546	LRRNT	Leucine rich repeat N-terminal domain	0.00082	26.1	1	62-84
1547	Allantoicase	Allantoicase repeat	7e-70	245.6	2	116-271:376-476
1547	DCX	Doublecortin	3.3e-17	70.6	1	611-671
1551	Peptidase_S9	Prolyl oligopeptidase family	1.7e-06	35.0	1	389-469
1551	DPPIV_N_term	Dipeptidyl peptidase IV (DPP IV) N-termi	1.4	-267.4	1	1-385
1551	DLH	Dienelactone hydrolase family	4.7	-96.4	1	355-597
1552	Orn_DAP_Ar_g_deC	Pyridoxal-dependent decarboxylase, C-	1.3e-05	23.1	1	84-202
1555	rvt	Reverse transcriptase	0.00015	7.9	1	505-718
1557	rvt	Reverse transcriptase	1.4e-69	244.6	1	247-520
1557	RbcX	RbcX protein	1.8	-79.3	1	1-108
1558	RhoGEF	RhoGEF domain	9.1e-26	99.0	1	971-1155
1558	PDZ	PDZ domain (Also known as DHR or GLGF)	4.2e-12	53.6	1	159-234
1558	PH	PH domain	0.081	19.5	1	1199-1312
1558	RGS	Regulator of G protein signaling domain	2.5	-34.3	1	465-577
1559	pkinase	Protein kinase domain	9e-11	20.4	1	151-443
1560	FH2	Formin Homology 2 Domain	4e-105	362.7	1	595-1038
1561	Keratin_B2	Keratin, high sulfur B2 protein	9.6	-83.6	1	136-235
1565	HMG_CoA_synt	Hydroxymethylglutaryl-coenzyme A synthas	6.5e-300	1009.7	1	50-582
1568	GCV_H	Glycine cleavage H-protein	4.3e-77	269.5	1	65-185



Table 4B  
489

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1570	DLH	Dienelactone hydrolase family	0.039	-69.3	1	168-431
1577	Peptidase_M1	Peptidase family M1	5.6e-74	259.2	1	42-399
1577	DUF59	Domain of unknown function DUF59	1.8	-24.4	1	188-265
1579	rrm	RNA recognition motif.	0.0082	22.8	1	823-891
1579	zf-C2H2	Zinc finger, C2H2 type	9.5	9.5	1	584-608
1580	serpin	Serpin (serine protease inhibitor)	2.1e-107	370.3	1	37-541
1581	CH	Calponin homology (CH) domain	0.012	13.5	1	84-181
1582	LRR	Leucine Rich Repeat	9.6e-38	138.8	12	70-93:94-117:118-141:142-165:166-189:190-213:214-237:238-261:262-285:286-310:311-335:336-359
1582	LRRCT	Leucine rich repeat C-terminal domain	5.6e-12	53.2	1	369-421
1582	ig	Immunoglobulin domain	1.9e-05	31.5	1	438-499
1582	LRRNT	Leucine rich repeat N-terminal domain	0.0099	22.5	1	28-68
1583	Nucleoplasmin	Nucleoplasmin	4.4	-94.1	1	131-199
1587	IBR	IBR domain	1.3e-13	58.7	1	237-298
1587	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.0053	20.0	1	168-213
1591	TPR	TPR Domain	4.1e-13	57.0	3	254-287:288-321:322-355
1592	PHD	PHD-finger	0.042	2.9	2	684-725:1117-1163
1592	DNA_ligase_ZBD	NAD-dependent DNA ligase C4 zinc finger	5.2	0.3	1	904-927
1593	Collagen	Collagen triple helix repeat (20 copies)	5.9	-64.0	1	302-360
1593	LIM	LIM domain	5.9	-18.2	1	841-870
1595	Peptidase_M1	Peptidase family M1	6.4e-44	159.3	1	7-279
1600	thyroglobulin_1	Thyroglobulin type-1 repeat	3.1e-06	34.1	1	38-118
1604	ig	Immunoglobulin domain	7	8.2	1	34-92
1605	PHD	PHD-finger	0.59	-7.8	1	585-644
1606	zf-CXXC	CXXC zinc finger	1.7e-21	84.9	1	50-96
1606	F-box	F-box domain	5e-06	33.4	2	545-602:713-762
1606	PHD	PHD-finger	0.062	1.3	1	207-271
1606	rubredoxin	Rubredoxin	3.6	-9.0	1	229-276
1611	L27	L27 domain	3.5	5.2	1	285-341
1612	fn3	Fibronectin type III domain	0.68	5.6	1	357-443
1613	rrm	RNA recognition motif.	3.8e-14	60.4	1	151-221
1614	SRCR	Scavenger receptor cysteine-rich domain	2.2e-26	101.1	1	52-145
1616	ldl_recept_a	Low-density lipoprotein receptor domain	0.12	8.2	1	68-114
1616	Tropomyosin	Tropomyosin	7.3	-118.6	1	216-441
1617	B56	Protein phosphatase 2A regulatory B subunit	3.3	-315.0	1	146-232

Table 4B  
490

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1618	S1	S1 RNA binding domain	6.7e-10	46.3	1	352-429
1618	HHH	Helix-hairpin-helix motif	3.8	12.3	1	173-202
1619	trypsin	Trypsin	5.9e-09	43.2	1	30-190
1623	FlgI	Flagellar P-ring protein	6.7e-19	76.2	1	46-257
1623	FlgH	Flagellar L-ring protein	1.5	-75.1	1	1-73
1624	Dihydroorotase	Dihydroorotase-like	9.9	-195.7	1	1-187
1625	Semialdehyde_dhC	Semialdehyde dehydrogenase, dimerisat	1.3e-54	194.8	1	137-338
1625	Semialdehyde_dh	Semialdehyde dehydrogenase, NAD bindi	2.3e-23	91.1	1	7-128
1626	FDX-ACB	Ferredoxin-fold anticodon binding domain	1e-43	158.6	1	83-176
1628	HTH_1	Bacterial regulatory helix-turn-helix	1.5e-13	58.4	1	7-66
1628	LysR_substrate	LysR substrate binding domain	2.5e-08	41.1	1	90-281
1628	Hpt	Hpt domain	7.8	-35.7	1	65-156
1631	PEP-utilizers_C	PEP-utilizing enzyme, TIM barrel doma	1.5e-10	-43.4	1	58-174
1632	OTCace	Aspartate/ornithine carbamoyltransferase, A	0.24	-67.2	1	432-546
1632	Replicase	Replicase family	2.9	-102.2	1	399-576
1634	ribonuclease_T2	Ribonuclease T2 family	0.00089	-47.2	1	31-165
1638	efhand	EF hand	5.9e-10	46.5	4	38-66:74-102:111-139:147-175
1639	PCI	PCI domain	0.013	19.5	1	165-246
1641	Cys_knot	Cystine-knot domain	0.011	-2.5	1	227-337
1641	DAN	DAN domain	0.28	-29.8	1	199-327
1643	Ribosomal_L23	Ribosomal protein L23	9.6e-13	55.8	1	197-276
1645	tsp_1	Thrombospondin type 1 domain	0.0026	21.0	5	420-476:480-534:541-601:607-663:669-718
1647	Arginosuc_synth	Arginosuccinate synthase	1.3e-07	-97.2	1	45-196
1651	vwa	von Willebrand factor type A domain	1.1e-33	125.4	1	89-326
1652	ig	Immunoglobulin domain	2.4e-09	44.5	2	34-117:220-296
1658	thioredo	Thioredoxin	6.9e-104	358.5	2	324-432:459-570
1658	glutaredoxin	Glutaredoxin	7.4	-4.2	1	482-544
1662	Hist_deacetyl	Histone deacetylase family	5.9e-61	215.9	1	857-1434
1666	wap	WAP-type (Whey Acidic Protein) 'four-disulfi	7.4e-15	62.8	1	72-116
1667	KRAB	KRAB box	1e-26	102.2	1	106-146
1672	ig	Immunoglobulin domain	3.5e-11	50.6	3	146-203:245-295:331-405
1674	Transposase_22	L1 transposable element	1.5e-51	184.7	1	6-281
1676	Ribosomal_S2	Ribosomal protein S2	0.19	-72.3	1	48-145
1677	Patatin	Patatin-like phospholipase	1.4e-36	135.0	1	132-301
1678	ig	Immunoglobulin domain	7.6	7.9	1	35-126
1678	Gemini_mov	Geminivirus putative	9.4	-42.6	1	51-131

Table 4B  
491

SEQ ID	Model	Description	E-value	Score	Repeats	Position
		movement protein				
1684	adh_short	short chain dehydrogenase	1.3e-32	121.8	1	2-244
1685	UPAR_LY6	u-PAR/Ly-6 domain	0.0039	12.2	1	131-213
1686	pkinase	Protein kinase domain	2e-43	157.7	1	553-900
1686	Huntingtin	Huntingtin	7.3	-163.8	1	520-893
1687	HRM	Hormone receptor domain	7.8e-05	29.5	1	82-134
1688	SH3	SH3 domain	5.4e-22	86.5	1	150-206
1690	ACAT	Sterol O-acyltransferase	6	-109.5	1	181-304
1693	LRR	Leucine Rich Repeat	4.3e-34	126.7	10	505-528:529-554:555-573:575-598:613-632:634-657:659-684:686-708:709-732:738-763
1694	SH3	SH3 domain	0.00089	19.6	1	19-73
1697	Metallothio_11	Metallothionein family 11	5.8	-11.1	1	36-81
1701	cystatin	Cystatin domain	0.01	11.8	1	32-125
1704	TPR	TPR Domain	9	6.5	1	122-155
1705	Ribosomal_S26e	Ribosomal protein S26e	2.1e-16	68.0	1	133-231
1706	ank	Ankyrin repeat	5.5e-30	113.1	3	381-413:414-446:450-482
1706	TPR	TPR Domain	1.6e-06	35.1	4	3-36:43-76:164-197:205-238
1706	LRR	Leucine Rich Repeat	0.0014	25.4	4	716-743:744-764:775-802:835-859
1706	Tropomodulin	Tropomodulin	2.6	-158.7	1	248-839
1708	zf-C2H2	Zinc finger, C2H2 type	3.6e-23	90.4	4	51-74:205-227:233-255:261-284
1708	zf-BED	BED zinc finger	7.7	-6.3	1	246-285
1709	Pep_M12B_propep	Reprolysin family propeptide	0.055	-20.1	1	120-202
1710	TIG	IPT/TIG domain	1.3e-100	347.7	10	620-705:708-788:793-876:883-988:1084-1169:1174-1254:1257-1336:1341-1423:1424-1511:1516-1602
1711	pp-binding	Phosphopantetheine attachment site	0.048	12.0	1	201-255
1711	HK	Hydroxyethylthiazole kinase family	2.9	-161.9	1	12-232
1711	ketoacyl-synt	Beta-ketoacyl synthase, N-terminal domain	4	-114.1	1	265-397
1712	IGFBP	Insulin-like growth factor binding protein	8.8e-26	99.1	1	40-99
1712	thyroglobulin_1	Thyroglobulin type-1 repeat	4.6e-05	29.5	1	264-323
1714	Paralemmmin	Paralemmmin	4.2e-26	100.1	1	101-439

Table 4B  
492

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1715	rvt	Reverse transcriptase	3.1e-08	40.8	1	2-147
1717	A2M_N	Alpha-2-macroglobulin family N-terminal r	0	1082.3	1	52-684
1717	A2M	Alpha-2-macroglobulin family	0	1184.5	1	776-1530
1717	TCTP	Translationally controlled tumor protein	0.0013	-47.0	1	1172-1249
1717	prenyltrans	Prenyltransferase and squalene oxidase re	7.6	3.0	1	1296-1332
1717	TSPN	Thrombospondin N-terminal -like domain	8.2	-57.8	1	34-202
1718	cadherin	Cadherin domain	9.7e-29	108.9	3	68-159;341-425;439-543
1721	MBOAT	MBOAT family	0.38	-121.1	1	3-173
1722	p450	Cytochrome P450	9.4e-48	172.1	1	86-527
1723	DUF270	Protein of unknown function, DUF270	0.12	-272.7	1	9-284
1725	ig	Immunoglobulin domain	0.75	16.2	1	34-116
1729	PA	PA domain	1.4e-29	111.7	1	521-626
1729	Glyco_hydro_47	Glycosyl hydrolase family 47	6.2e-13	28.4	1	99-364
1731	7tm_1	7 transmembrane receptor (rhodopsin family)	3.1e-62	220.2	1	50-293
1731	7tm_5	7TM chemoreceptor	1.7	-157.3	1	31-289
1731	7tm_3	7 transmembrane receptor	6.1	-155.5	1	39-302
1732	ig	Immunoglobulin domain	9.3e-16	65.8	2	49-125;169-240
1732	Adeno_E3_CR1	Adenovirus E3 region protein CR1	8.5	-24.2	1	171-251
1734	p450	Cytochrome P450	0.0069	-119.1	1	51-492
1735	profilin	Profilin	5.1e-06	33.3	1	955-1071
1735	Transposase_22	L1 transposable element	0.00026	-102.7	1	731-1016
1735	KE2	KE2 family protein	7.8	-47.5	1	765-838
1739	IL3	Interleukin-3	8.7	-38.1	1	121-224
1742	PAS	PAS domain	0.024	21.2	1	348-413
1743	transmembrane4	Tetraspanin family	0.87	-77.9	1	5-194
1744	Dynein_heavy	Dynein heavy chain	2e-253	855.3	1	4-781
1745	SCP	SCP-like extracellular protein	9.8e-36	132.1	1	4-180
1745	granulin	Granulin	0.32	-8.1	1	173-206
1749	sugar_tr	Sugar (and other) transporter	0.26	-152.9	1	260-692
1749	oxidored_q1	NADH-Ubiquinone/plastoquinone	1.6	-155.3	1	391-674
1749	FecCD	FecCD transport family	4.7	-214.9	1	473-686
1749	Herpes_HEPA	Herpesvirus DNA helicase/primase compl	5.4	-457.3	1	77-605
1749	PsbT	Photosystem II reaction centre T prote	5.9	5.1	1	545-572
1749	Bac_rhodopsin	Bacteriorhodopsin	6.8	-137.3	1	295-496
1749	C4dic_mal_transporter	C4-dicarboxylate transporter/malic aci	9.9	-183.4	1	265-477

Table 4B  
493

SEQ ID	Model	Description	E-value	Score	Repeats	Position
1750	ATP-synt_A	ATP synthase A chain	8.1	-80.8	1	837-958
1751	SDF	Sodium:dicarboxylate symporter family	2.7	-254.2	1	275-426
1753	ig	Immunoglobulin domain	9.9e-19	75.7	5	49-124:162-219:261-310:345-400:432-496
1753	fn3	Fibronectin type III domain	0.00052	26.8	1	514-606
1755	sugar_tr	Sugar (and other) transporter	0.13	-144.2	1	1-366
1755	PUCC	PUCC protein	2.3	-267.4	1	21-314
1755	Nuc_H_symport	Nucleoside H <sup>+</sup> symporter	4.4	-267.2	1	1-356
1755	7tm_5	7TM chemoreceptor	4.8	-165.5	1	59-281
1755	Competence	Competence protein	5.1	-101.6	1	16-209
1755	Ammonium_transp	Ammonium Transporter Family	5.4	-242.0	1	67-380
1755	BenE	Benzoate membrane transport protein	8.1	-335.2	1	69-351
1757	Exo_endo_phos	Endonuclease/Exonuclease/phosphatase fa	1.3e-20	81.9	1	367-788
1758	K_tetra	K <sup>+</sup> channel tetramerisation domain	2.9	-38.6	1	114-214
1759	HK	Hydroxyethylthiazole kinase family	3	-162.1	1	303-502
1760	Peptidase_C20	Type IV leader peptidase family	6.6	-187.3	1	29-187
1762	SSF	Sodium:solute symporter family	4.2e-256	864.2	1	39-477
1762	MVIN	Virulence factor MVIN	7.1	-250.2	1	68-548
1762	PNTB	NAD(P) transhydrogenase beta subunit	8.1	-394.0	1	11-300
1762	7tm_5	7TM chemoreceptor	9.1	-170.5	1	75-352
1762	oxidored_q1	NADH-Ubiquinone/plastoquinone (complex I)	9.1	-170.9	1	145-463
1763	LRR	Leucine Rich Repeat	2e-37	137.7	11	61-84:85-108:109-132:133-156:157-180:181-204:205-228:253-276:277-300:301-324:349-371
1763	LRRCT	Leucine rich repeat C-terminal domain	1.9e-07	38.2	1	358-411
1763	ig	Immunoglobulin domain	8.5e-05	29.4	1	428-486
1763	LRRNT	Leucine rich repeat N-terminal domain	0.00023	27.9	1	30-59
1764	Reprolysin	Reprolysin (M12B) family zinc metalloprote	1e-21	85.6	1	156-359
1764	tsp_1	Thrombospondin type 1 domain	2.5e-06	34.5	1	422-491
1769	C2	C2 domain	2.5e-66	233.8	2	249-337:383-471
1771	cadherin	Cadherin domain	2.7e-79	276.8	5	140-235:249-343:359-451:465-556:607-697
1774	transmembran	Tetraspanin family	6.4e-59	209.2	1	458-702

Table 4B

494

SEQ ID	Model	Description	E-value	Score	Repeats	Position
	e4					
1775	7tm_1	7 transmembrane receptor (rhodopsin family)	7.9e-05	-5.9	1	34-333
1775	Srg	C.elegans Srg family integral membrane prote	9.3	-223.5	1	111-306
1778	synaptobrevin	Synaptobrevin	4.1e-32	120.1	1	110-198
1779	trefoil	Trefoil (P-type) domain	3.9	-11.1	1	229-254
1780	MHC_I	Class I Histocompatibility antigen, domains	2.6e-89	310.1	1	25-199
1780	ig	Immunoglobulin domain	3.3e-06	34.0	1	194-259
1782	ABC_tran	ABC transporter	2.4e-19	77.7	1	448-626
1783	SNF	Sodium:neurotransmitter symporter famil	0	1189.5	1	203-747
1783	Herpes_glycop	Herpesvirus glycoprotein M	7.2	-262.5	1	460-801
1783	UPF0118	Domain of unknown function DUF20	8	-131.6	1	240-582
1783	Cyto_ox_2	Cytochrome oxidase subunit II	9	-243.9	1	467-696
1819	Y_phosphatase	Protein-tyrosine phosphatase	1.2e-103	357.8	1	1277-1494
1819	fn3	Fibronectin type III domain	1.2e-51	185.1	7	461-541:552-634:641-719:730-812:819-900:908-990:998-1087
1821	7tm_1	7 transmembrane receptor (rhodopsin family)	4.4	-116.1	1	206-413
1822	Metallophos	Calcineurin-like phosphoesterase	1.5	2.4	1	66-252

Table 5

SEQ ID No.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
913	1ae7		472	598	5.1e-25	0.14	-0.09		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE PLA2, PHOSPHATIDE SN-2 ACYLHYDROLASE; HYDROLASE, PHOSPHOLIPASE A2, LIPID DEGRADATION, PRESYNAPTIC 2 NEUROTOXIN, VENOM
913	1aut	L	468	569	8.5e-19	0.17	0.03		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO- MAL; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
913	1aut	L	553	637	3.4e-15	0.01	-0.13		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO- MAL; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
913	1aut	L	784	842	5.1e-16	0.04	0.54		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO- MAL; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
913	1bk9		470	597	3.4e-26	0.32	-0.18		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE HYDROLASE, PHOSPHOLIPASE A2, PLATELET AGGREGATION INHIBITOR, 2 PBFB HEADER MODRES
913	1bk9		510	637	1.4e-25	0.17	-0.15		PHOSPHOLIPASE A2;	HYDROLASE HYDROLASE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: NULL;	PHOSPHOLIPASE A2, PLATELET AGGREGATION INHIBITOR, 2 PBPB HEADER MODRES
913	1dan	L	430	520	7e-11	-0.24	0.18		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
913	1dan	L	455	565	8.5e-21	-0.06	0.00		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
913	1dan	L	482	602	3.4e-26	0.06	0.34		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
913	1dan	L	527	637	6.8e-21	-0.23	0.60		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
913	1dan	L	782	897	3.4e-19	-0.10	0.07		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
913	1dqb	A	468	558	3.4e-21	0.28	0.15		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
913	1dqb	A	796	876	1.4e-17	0.03	-0.09		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
913	1dva	L	430	520	7e-11	0.08	0.39		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
913	1dva	L	468	565	3.4e-19	0.05	0.25		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
913	1dva	L	507	602	3.4e-19	0.12	0.53		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	
913	1dva	L	551	637	1.7e-16	0.18	0.81		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
913	1dva	L	784	852	5.1e-14	-0.17	0.05		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
913	1dva	L	795	897	3.4e-17	-0.13	0.07		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
913	1dx5	I	395	508	1.4e-15	-0.16	0.11		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN, EGR-CMK SERINE PROTEINASE, EGF-LIKE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
913	1dx5	I	454	553	1.4e-17	0.24	0.18		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
913	1dx5	I	470	593	6.8e-29	0.29	0.86		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
913	1dx5	I	721	836	2.8e-14	-0.08	0.10		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
913	1dx5	I	784	886	8.5e-18	-0.12	0.07		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE

500

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
913	1emm		430	509	2.8e-15	0.40	0.86		CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H; FIBRILLIN; CHAIN: NULL;	PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
913	1emm		552	611	3.4e-21	-0.45	0.74		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
913	1emm		683	752	5.6e-10	0.02	-0.05		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
913	1emm		795	857	1e-19	-0.23	0.83		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
913	1emm		838	908	1.1e-09	-0.55	0.12		FIBRILLIN; CHAIN: NULL;	FRAGMENT, MATRIX PROTEIN
913	1ext	A	453	624	3.4e-31	0.14	0.36		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
913	1ext	A	70	220	1.3e-08	0.25	-0.19		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALING PROTEIN
913	1ext	A	785	888	8.5e-21	0.34	-0.11		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALING PROTEIN
913	1fak	L	430	520	7e-11	0.04	0.21		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
913	1fak	L	468	565	6.8e-20	0.27	0.39		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										(SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
913	1fak	L	506	602	1.7e-20	0.18	0.29		BLOOD COAGULATION FACTOR VIIA; CHAIN: I; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
913	1fak	L	552	637	8.5e-16	-0.15	0.19		BLOOD COAGULATION FACTOR VIIA; CHAIN: I; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
913	1fak	L	795	897	6.8e-17	-0.07	0.30		BLOOD COAGULATION FACTOR VIIA; CHAIN: I; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
913	1h4u	A	227	446	3.4e-55	0.49	0.95		NIDOGEN-1; CHAIN: A;	EXTRACELLULAR MATRIX PROTEIN EXTRACELLULAR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
913	1hj7	A	114	187	2.8e-09	0.32	-0.20		LDL RECEPTOR; CHAIN: A;	MATRIX PROTEIN
										CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
913	1hj7	A	595	720	1.4e-12	-0.65	0.10		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
										CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
913	1hj7	A	784	837	1.4e-18	0.66	1.00		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
										CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
913	1hj7	A	800	882	1.7e-21	0.19	0.70		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
										CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
913	1hj7	A	840	928	1.4e-08	-0.22	0.00		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
										CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
913	1igr	A	96	253	1.3e-08	0.05	-0.19		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
										HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
913	1ijq	A	344	508	6.8e-15	0.23	0.06		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A, B;	LIPID TRANSPORT LDL RECEPTOR, BETA-PROPELLER
										LIPID TRANSPORT LDL RECEPTOR, BETA-PROPELLER
913	1kdo		129	286	1.1e-09	0.24	-0.20		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
										GLYCOPROTEIN GLYCOPROTEIN
913	1kdo		172	325	1.3e-08	0.02	-0.20		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
										GLYCOPROTEIN GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
913	1klo		419	554	9.8e-11	0.11	-0.15		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
913	1klo		474	633	1.7e-21	0.31	-0.07		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
913	1klo		70	231	1.4e-11	0.16	-0.18		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
913	1pfx	L	454	578	1.7e-18	0.10	-0.03		FACTOR IXA; CHAIN: C; L <sub>2</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
913	1pfx	L	478	610	8.5e-21	0.02	0.11		FACTOR IXA; CHAIN: C; L <sub>2</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
913	1pfx	L	522	637	3.4e-19	0.07	-0.03		FACTOR IXA; CHAIN: C; L <sub>2</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
913	1pfx	L	787	888	5.1e-17	-0.06	0.00		FACTOR IXA; CHAIN: C; L <sub>2</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
913	1qfk	L	784	859	1.7e-15	0.18	0.10		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
913	1vap	A	472	597	3.4e-24	0.14	-0.14		PHOSPHOLIPASE A2; CHAIN: A, B;	LIPID DEGRADATION PHOSPHOLIPASE A2, LIPID DEGRADATION, HYDROLASE
913	1vap	A	511	637	1.7e-23	0.19	-0.17		PHOSPHOLIPASE A2; CHAIN: A, B;	LIPID DEGRADATION PHOSPHOLIPASE A2, LIPID DEGRADATION, HYDROLASE
913	1vmo	A	121	235	3.4e-12	0.70	-0.08		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I1VMO3	
913	1xka	L	558	637	5.1e-14	0.27	0.81		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2
913	1xka	L	784	842	3.4e-14	0.11	0.68		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2
913	9wga	A	102	272	1.3e-14	0.07	-0.18		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	GROWTH FACTOR LIKE DOMAIN
913	9wga	A	353	493	4.2e-10	0.34	-0.18		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
914	1dpt	A	2	111	1.1e-33			130.55	(ISOLECTIN 2) 9WGA 3 D-DOPACHROME TAUTOMERASE; CHAIN: A, B, C;	CYTOKINE CYTOKINE, GROWTH FACTOR, TAUTOMERASE
914	1dpt	A	2	115	2.8e-33			133.05	D-DOPACHROME TAUTOMERASE; CHAIN: A, B, C;	CYTOKINE CYTOKINE, GROWTH FACTOR, TAUTOMERASE
914	1dpt	A	2	95	2.8e-33	0.98	1.00		D-DOPACHROME TAUTOMERASE; CHAIN: A, B, C;	CYTOKINE CYTOKINE, GROWTH FACTOR, TAUTOMERASE
914	1f5m		2	109	9.8e-37			60.34	MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: NULL;	CYTOKINE MACROPHAGE, INFLAMMATORY RESPONSE, CYTOKINE
914	1f5m		2	95	1.4e-36	0.69	1.00		MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: NULL;	CYTOKINE MACROPHAGE, INFLAMMATORY RESPONSE, CYTOKINE
914	1gd0	A	2	111	5.6e-37			60.28	MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	IMMUNE SYSTEM MIF, MIF, MACROPHAGE MIGRATION INHIBITORY FACTOR,
914	1gd0	A	2	123	1.1e-36			69.95	MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	IMMUNE SYSTEM MIF, MIF, MACROPHAGE MIGRATION INHIBITORY FACTOR,
914	1gd0	A	2	95	1.1e-36	0.69	1.00		MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	IMMUNE SYSTEM MIF, MIF, MACROPHAGE MIGRATION INHIBITORY FACTOR,
914	1mfi	A	2	120	4.2e-37			67.20	MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	CYTOKINE PHENYLPIRUVATE TAUTOMERASE, CYTOKINE, MACROPHAGE, INFLAMMATORY RESPONSE, TAUTOMERASE

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
914	1mfi	A	2	95	4.2e-37	0.85	1.00		MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	CYTOKINE PHENYLPIRUVATE TAUTOMERASE; CYTOKINE, MACROPHAGE, INFLAMMATORY RESPONSE, TAUTOMERASE
915	1dpt	A	2	131	2.8e-42	0.69	1.00		D-DOPACHROME TAUTOMERASE; CHAIN: A, B, C;	CYTOKINE CYTOKINE, GROWTH FACTOR, TAUTOMERASE
915	1dpt	A	2	131	2.8e-42			163.64	D-DOPACHROME TAUTOMERASE; CHAIN: A, B, C;	CYTOKINE CYTOKINE, GROWTH FACTOR, TAUTOMERASE
915	1gd0	A	2	127	1.4e-44	0.60	1.00		MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	IMMUNE SYSTEM MIF, MIF, MACROPHAGE MIGRATION INHIBITORY FACTOR,
915	1gd0	A	2	130	1.4e-44			63.03	MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	IMMUNE SYSTEM MIF, MIF, MACROPHAGE MIGRATION INHIBITORY FACTOR,
915	1mfi	A	2	127	9.8e-46	0.72	0.98		MACROPHAGE MIGRATION INHIBITORY FACTOR; CHAIN: A, B, C;	CYTOKINE PHENYLPIRUVATE TAUTOMERASE; CYTOKINE, MACROPHAGE, INFLAMMATORY RESPONSE, TAUTOMERASE
916	1a0h	A	284	467	1.7e-29	-0.30	0.39		MEIZOTROMBIN; CHAIN: A, B, D, E; D-PHE-PRO-ARG; CHAIN: C, F;	COMPLEX (SERINE PROTEASE/INHIBITOR) DESF1; PPACK; SERINE PROTEASE, COAGULATION, THROMBIN, PROTHROMBIN, 2
916	1a0j	A	11	248	4.2e-96			190.46	TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
916	1a0j	A	445	682	2.8e-96	0.93	1.00		TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQ:COL D score	Compound	PDB annotation
916	1a0j	A	445	682	2.8e-96			193.09	TRYPSIN; CHAIN: A, B, C, D;	HYDROLASE SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
916	1a5i	A	1	247	5.6e-83			214.60	PLASMINOGEN ACTIVATOR; CHAIN: A; GLU-GLY-ARG CHLOROMETHYL KETONE; CHAIN: I;	COMPLEX (SERINE PROTEASE/INHIBITOR) (DELTAPEK)DSPALPHA1; EGRCMK; SERINE PROTEASE, FIBRINOLYTIC ENZYMES, PLASMINOGEN 2 ACTIVATORS
916	1a5i	A	430	681	1.2e-85			219.75	PLASMINOGEN ACTIVATOR; CHAIN: A; GLU-GLY-ARG CHLOROMETHYL KETONE; CHAIN: I;	COMPLEX (SERINE PROTEASE/INHIBITOR) (DELTAPEK)DSPALPHA1; EGRCMK; SERINE PROTEASE, FIBRINOLYTIC ENZYMES, PLASMINOGEN 2 ACTIVATORS
916	1aut	C	11	246	1.4e-78			182.91	ACTIVATED PROTEIN C; CHAIN: C, I; D-PHE-PRO- MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
916	1aut	C	445	680	5.1e-79			185.86	ACTIVATED PROTEIN C; CHAIN: C, I; D-PHE-PRO- MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
916	1aut	L	159	252	8.4e-15	0.32	-0.03		ACTIVATED PROTEIN C; CHAIN: C, I; D-PHE-PRO- MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
										HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
916	1b2i	A	284	367	1.4e-28	-0.06	1.00		PLASMINOGEN; CHAIN: A;	HYDROLASE SERINE PROTEASE, FIBRINOLYSIS, LYSINE-BINDING DOMAIN, 2 PLASMINOGEN, KRINGLE 2, HYDROLASE
916	1b2i	A	76	153	4.2e-21	0.20	-0.14		PLASMINOGEN; CHAIN: A;	HYDROLASE SERINE PROTEASE, FIBRINOLYSIS, LYSINE-BINDING DOMAIN, 2 PLASMINOGEN, KRINGLE 2, HYDROLASE
916	1b2i	A	7	93	2.8e-28			70.19	PLASMINOGEN; CHAIN: A;	HYDROLASE SERINE PROTEASE, FIBRINOLYSIS, LYSINE-BINDING DOMAIN, 2 PLASMINOGEN, KRINGLE 2, HYDROLASE
916	1b29		162	201	1.2e-13	0.80	1.00		FACTOR VII; CHAIN: NULL;	BLOOD COAGULATION BLOOD COAGULATION, EGF, HYDROLASE, SERINE PROTEASE
916	1bru	P	11	248	4.2e-87			180.84	ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
916	1bru	P	445	682	1.4e-88	0.96	1.00		ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
916	1bru	P	445	682	1.4e-88			183.63	ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
916	1c5y	B	11	248	2.8e-85			214.59	UROKINASE-TYPE PLASMINOGEN ACTIVATOR; CHAIN: A;	BLOOD CLOTTING SELECTIVE, S1 SITE INHIBITOR, STRUCTURE- BASED DRUG DESIGN, 2 UROKINASE, TRYPSIN, THROMBIN
916	1c5y	B	445	682	8.5e-86			217.96	UROKINASE-TYPE PLASMINOGEN	BLOOD CLOTTING SELECTIVE, S1 SITE INHIBITOR, STRUCTURE-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ACTIVATOR; CHAIN: A; UROKINASE-TYPE PLASMINOGEN ACTIVATOR; CHAIN: B; PLASMINOGEN; ICEA 7 CHAIN: A, B, ICEA 8	BASED DRUG DESIGN, 2 UROKINASE, TRYPSIN, THROMBIN
916	1cea	A	10	92	1.4e-26			65.46		SERINE PROTEASE K1PG, ICEA 10
916	1cxw	A	102	151	3.4e-20	0.72	1.00		HUMAN MATRIX METALLOPROTEINASE 2; CHAIN: A; HUMAN MATRIX METALLOPROTEINASE 2; CHAIN: A; HYDROLASE COL-2; BETA SHEET, ALPHA HELIX, HYDROLASE	HYDROLASE COL-2; BETA SHEET, ALPHA HELIX, HYDROLASE
916	1cxw	A	37	95	7e-24			60.11		HYDROLASE COL-2; BETA SHEET, ALPHA HELIX, HYDROLASE
916	1dan	H	11	251	1.1e-77			170.95	BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D- PHE-PHE-ARG- CHLOROMETHYLKETO NE (DFRCMK) WITH CHAIN: C; BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D- PHE-PHE-ARG- CHLOROMETHYLKETO NE (DFRCMK) WITH CHAIN: C; BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
916	1dan	L	133	251	3.4e-18	0.16	-0.08			BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
916	1dan	L	162	250	1.4e-27	0.23	0.68		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D- PHE-PHE-ARG- CHLOROMETHYLKETO NE (DFRCMK) WITH CHAIN: C; BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
916	1ddj	A	1	248	2.8e-89			176.67	CHAIN: C; PLASMINOGEN; CHAIN: A, B, C, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC DOMAIN
916	1ddj	A	428	682	7e-95	0.87	1.00		PLASMINOGEN; CHAIN: A, B, C, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC DOMAIN
916	1ddj	A	429	682	7e-95			187.88	PLASMINOGEN; CHAIN: A, B, C, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC DOMAIN
916	1dlk	B	11	248	4.2e-82			179.88	DELTA-CHYMOTRYPSIN; CHAIN: A, C; DELTA-CHYMOTRYPSIN; CHAIN: B, D;	HYDROLASE DELTA-CHYMOTRYPSIN, PEPTIDIC INHIBITOR, CHLOROMETHYL KETONE
916	1dlk	B	445	682	1.3e-83			183.37	DELTA-CHYMOTRYPSIN; CHAIN: A, C; DELTA-CHYMOTRYPSIN; CHAIN: B, D;	HYDROLASE DELTA-CHYMOTRYPSIN, PEPTIDIC INHIBITOR, CHLOROMETHYL KETONE
916	1dva	L	162	250	1.4e-27	0.39	0.47		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
916	1dva	L	4	106	4.2e-19			64.99	DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQPOL D score	Compound	PDB annotation
916	1dx5	I	5	126	2.8e-18			54.24	THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H; Y;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
916	1e88	A	1	125	5.6e-34			56.65	FIBRONECTIN; CHAIN: A;	EXTRACELLULAR MATRIX GLYCOPROTEIN
916	1e88	A	97	207	6.8e-21	0.55	0.33		FIBRONECTIN; CHAIN: A;	EXTRACELLULAR MATRIX GLYCOPROTEIN
916	1ekb	B	11	246	9.8e-83			190.18	ENTEROPEPTIDASE; CHAIN: A; ENTEROPEPTIDASE; CHAIN: B; VAL-ASP-ASP-ASP-LYS PEPTIDE; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPSINOGEN ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR
916	1ekb	B	445	680	1.4e-82			193.02	ENTEROPEPTIDASE; CHAIN: A; ENTEROPEPTIDASE; CHAIN: B; VAL-ASP-ASP-ASP-LYS PEPTIDE; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPSINOGEN ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQIDOL D score	Compound	PDB annotation
916	1elt		11	245	2.8e-81			175.26	ELASTASE; 1ELT 4 CHAIN: NULL; 1ELT 5	SERINE PROTEINASE
916	1elt		445	679	2.8e-81			178.10	ELASTASE; 1ELT 4 CHAIN: NULL; 1ELT 5	SERINE PROTEINASE
916	1ext	A	146	273	5.1e-13	0.33	-0.13		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
916	1f7z	A	11	248	8.4e-92			170.59	TRYPSIN II, ANIONIC; CHAIN: A; PANCREATIC TRYPSIN INHIBITOR; CHAIN: I;	HYDROLASE/HYDROLASE INHIBITOR BPTI SERINE PROTEASE, TRYPSIN PRECURSOR
916	1f7z	A	445	682	2.8e-92	0.92	1.00		TRYPSIN II, ANIONIC; CHAIN: A; PANCREATIC TRYPSIN INHIBITOR; CHAIN: I;	HYDROLASE/HYDROLASE INHIBITOR BPTI SERINE PROTEASE, TRYPSIN PRECURSOR
916	1fak	L	162	250	1.4e-27	0.35	0.66		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX/SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
916	1fak	L	4	106	4.2e-19			59.97	BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX/SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
916	1ftw	A	11	249	2.8e-81			169.75	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
916	1ftw	A	445	689	4.2e-81			177.29	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
916	1ftz	A	11	249	2.8e-81			171.91	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
916	1ftz	A	445	689	5.6e-82			177.71	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
916	1fti	A	11	248	2.8e-97			186.43	TRYPSIN; CHAIN: A;	HYDROLASE SERINE PROTEASE, HYDROLASE
916	1fti	A	445	682	5.6e-98	1.03	1.00		TRYPSIN; CHAIN: A;	HYDROLASE SERINE PROTEASE, HYDROLASE
916	1fti	A	445	682	5.6e-98			189.59	TRYPSIN; CHAIN: A;	HYDROLASE SERINE PROTEASE, HYDROLASE
916	1fxy	A	11	249	4.2e-86			178.61	COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE (PPACK) WITH CHAIN: L;	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
916	1fxy	A	445	682	5.6e-87	1.02	1.00		COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									NE (PPACK) WITH CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR)
916	1fxv	A	445	683	5.6e-87			181.30	COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE (PPACK) WITH CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
916	1gct	A	1	248	4.2e-80			174.02	HYDROLASE (SERINE PROTEINASE) GAMMA-CHYMOTRYPSIN *A (E.C.3.4.21.1) (\$P*H 7.0) 1GCT 3	
916	1gct	A	431	682	2.8e-83			176.84	HYDROLASE (SERINE PROTEINASE) GAMMA-CHYMOTRYPSIN *A (E.C.3.4.21.1) (\$P*H 7.0) 1GCT 3	
916	1hq8	A	49	163	2.8e-29			82.88	NKG2-D; CHAIN: A;	APOPTOSIS HOMODIMER, CIS-PROLINE
916	1i5k	A	10	92	1.4e-28			68.60	PLASMINOGEN; CHAIN: A, B; M PROTEIN; CHAIN: C, D;	BLOOD CLOTTING HUMAN PLASMINOGEN KRINGLE-2, KRINGLES, VEK-30
916	1i5k	A	78	153	8.4e-22	0.12	-0.18		PLASMINOGEN; CHAIN: A, B; M PROTEIN; CHAIN: C, D;	BLOOD CLOTTING HUMAN PLASMINOGEN KRINGLE-2, KRINGLES, VEK-30
916	1i71	A	10	96	2.8e-30			73.08	APOLIPROTEIN(A); CHAIN: A;	HYDROLASE APO(A); ALIPOPTEIN(A), KRINGLE, PROTEIN-LIGAND INTERACTION, 2 LYSINE BINDING, CRYSTAL STRUCTURE
916	1i71	A	77	154	8.4e-21	0.26	-0.13		APOLIPROTEIN(A); CHAIN: A;	HYDROLASE APO(A); ALIPOPTEIN(A), KRINGLE,

516

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
916	1j7m	A	105	150	5.1e-19	0.93	1.00		MATRIX METALLOPROTEINASE 2; CHAIN: A;	PROTEIN-LIGAND INTERACTION, 2 LYSINE BINDING, CRYSTAL STRUCTURE
916	1kdu		10	94	2.8e-31			79.24	PLASMINOGEN ACTIVATION PLASMINOGEN ACTIVATOR (UROKINASE-TYPE, KRINGLE DOMAIN) IKDU 3 (U-PA K) (NMR, MINIMIZED AVERAGE STRUCTURE) IKDU 4	
916	1kdu		285	369	7e-30	0.47	1.00		PLASMINOGEN ACTIVATION PLASMINOGEN ACTIVATOR (UROKINASE-TYPE, KRINGLE DOMAIN) IKDU 3 (U-PA K) (NMR, MINIMIZED AVERAGE STRUCTURE) IKDU 4	
916	1kdu		286	367	1.7e-32	0.43	1.00		PLASMINOGEN ACTIVATION PLASMINOGEN ACTIVATOR (UROKINASE-TYPE, KRINGLE DOMAIN) IKDU 3 (U-PA K) (NMR, MINIMIZED AVERAGE STRUCTURE) IKDU 4	
916	1klo		146	293	1e-14	0.25	-0.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
916	1kcn		10	92	2.8e-32			76.64	PLASMINOGEN; CHAIN:	SERINE PROTEASE KRINGLE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									NULL;	BLOOD, PLASMINOGEN, SERINE PROTEASE
916	1km		285	367	2.8e-28	0.21	1.00		PLASMINOGEN; CHAIN: NULL;	SERINE PROTEASE KRINGLE, BLOOD, PLASMINOGEN, SERINE PROTEASE
916	1km		77	153	9.8e-23	0.21	-0.09		PLASMINOGEN; CHAIN: NULL;	SERINE PROTEASE KRINGLE, BLOOD, PLASMINOGEN, SERINE PROTEASE
916	1pnm	A	11	246	5.6e-80			172.80	NEUROPSIN; CHAIN: A, B;	SERINE PROTEINASE SERINE PROTEINASE, GLYCOPROTEIN
916	1pnm	A	445	680	5.6e-80			175.97	NEUROPSIN; CHAIN: A, B;	SERINE PROTEINASE SERINE PROTEINASE, GLYCOPROTEIN
916	1pdc		104	148	1.2e-21	0.52	1.00		COLLAGEN-BINDING TYPE II DOMAIN SEMINAL FLUID PROTEIN PDC-109 (DOMAIN B) 1PDC 3 (NMR, BEST STRUCTURE) 1PDC 4	
916	1pdc		47	91	5.6e-18			54.17	COLLAGEN-BINDING TYPE II DOMAIN SEMINAL FLUID PROTEIN PDC-109 (DOMAIN B) 1PDC 3 (NMR, BEST STRUCTURE) 1PDC 4	
916	1pfx	C	11	248	1.1e-83			180.03	FACTOR IXA; CHAIN: C, L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN

518

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQ/ROL D score	Compound	PDB annotation
916	1pfx	C	445	682	2.8e-83			181.67	FACTOR IXA; CHAIN: C; L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
916	1pfx	L	122	267	1e-18	0.17	-0.13		FACTOR IXA; CHAIN: C; L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
916	1pfx	L	162	259	4.2e-23	0.39	-0.12		FACTOR IXA; CHAIN: C; L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
916	1pml	A	286	367	8.5e-32	0.72	1.00		HYDROLASE(SERINE PROTEASE) TISSUE PLASMINOGEN ACTIVATOR KRINGLE 2 (E.C.3.4.21.68) 1PML 3	
916	1pml	A	9	94	4.2e-32			101.06	HYDROLASE(SERINE PROTEASE) TISSUE PLASMINOGEN ACTIVATOR KRINGLE 2 (E.C.3.4.21.68) 1PML 3	
916	1pml	C	286	367	1.2e-31	0.71	1.00		HYDROLASE(SERINE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEASE) TISSUE PLASMINOGEN ACTIVATOR KRINGLE 2 (E.C.3.4.21.68) 1PML 3	
916	1pml	C	9	93	4.2e-32			101.45	HYDROLASE/SERINE PROTEASE) TISSUE PLASMINOGEN ACTIVATOR KRINGLE 2 (E.C.3.4.21.68) 1PML 3	
916	1pp2	L	162	284	1.7e-23	0.02	-0.18		HYDROLASE CALCIUM-FREE PHOSPHOLIPASE A=2= (E.C.3.1.1.4) 1PP2 4	
916	1pyt	D	1	247	5.6e-81			182.06	PROCARBOXYPEPTIDAS E A; CHAIN: A, B; PROPROTEINASE E; CHAIN: C; CHYMOTRYPSINOGEN C; CHAIN: D;	TERNARY COMPLEX (ZYMOGEN) TC, PCPA-TG; TERNARY COMPLEX (ZYMOGEN), SERINE PROTEINASE, C-TERMINAL 2 PEPTIDASE
916	1pyt	D	431	681	8.4e-83			187.81	PROCARBOXYPEPTIDAS E A; CHAIN: A, B; PROPROTEINASE E; CHAIN: C; CHYMOTRYPSINOGEN C; CHAIN: D;	TERNARY COMPLEX (ZYMOGEN) TC, PCPA-TG; TERNARY COMPLEX (ZYMOGEN), SERINE PROTEINASE, C-TERMINAL 2 PEPTIDASE
916	1qfk	L	10	109	2.8e-18			58.07	COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
916	1qfk	L	163	250	4.2e-27	0.17	0.12		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C; FIBRONECTIN; CHAIN: A;	
916	1qo6	A	102	152	3.4e-21	1.10	1.00			CELL ADHESION PROTEIN FIBRONECTIN MODULE PAIR, GELATIN-BINDING
916	1rfh	A	11	248	2.8e-83			181.07	COAGULATION FACTOR IX; CHAIN: A; COAGULATION FACTOR IX; CHAIN: B;	COAGULATION FACTOR SERINE PROTEINASE, BLOOD COAGULATION, COAGULATION FACTOR
916	1rfh	A	445	682	1.1e-82			183.25	COAGULATION FACTOR IX; CHAIN: A; COAGULATION FACTOR IX; CHAIN: B;	COAGULATION FACTOR SERINE PROTEINASE, BLOOD COAGULATION, COAGULATION FACTOR
916	1rtf	B	11	247	1.4e-85			229.36	TWO CHAIN TISSUE PLASMINOGEN ACTIVATOR; CHAIN: A, B;	SERINE PROTEASE (TC)-T-PA; SERINE PROTEASE, FIBRINOLYTIC ENZYMES
916	1rtf	B	445	681	1e-84			231.89	TWO CHAIN TISSUE PLASMINOGEN ACTIVATOR; CHAIN: A, B;	SERINE PROTEASE (TC)-T-PA; SERINE PROTEASE, FIBRINOLYTIC ENZYMES
916	1slw	B	11	248	2.8e-92			174.20	ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
916	1slw	B	445	682	8.4e-93	1.01	1.00		ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
916	1slw	B	445	682	8.4e-93			176.17	ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
										COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
916	1tpg		202	285	1.4e-24	0.61	1.00		T-PLASMINOGEN ACTIVATOR F1-G; 1TPG 7 CHAIN: NULL; 1TPG 8	PLASMINOGEN ACTIVATION
916	1tpg		45	132	9.8e-19			75.25	T-PLASMINOGEN ACTIVATOR F1-G; 1TPG 7 CHAIN: NULL; 1TPG 8	PLASMINOGEN ACTIVATION
916	1tm	A	11	249	2.8e-94			185.24	HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR 1TRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) 1TRN 4 HUMAN TRYPSIN, DFP INHIBITED 1TRN 6	
916	1tm	A	445	682	9.8e-95	0.85	1.00		HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR 1TRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) 1TRN 4 HUMAN TRYPSIN, DFP INHIBITED 1TRN 6	

522

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL ID score	Compound	PDB annotation
									RIDATE (DFP) ITRN 4 HUMAN TRYPSIN, DFP INHIBITED ITRN 6	
916	1tm	A	445	683	9.8e-95			188.29	HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR ITRN 3 DIISOPROPYL- FLUOROPHOSPHOFLUO RIDATE (DFP) ITRN 4 HUMAN TRYPSIN, DFP INHIBITED ITRN 6	
916	1urk		1	96	1.3e-34			70.09	PLASMINOGEN ACTIVATION PLASMINOGEN ACTIVATOR (UROKINASE-TYPE) (AMINO TERMINAL FRAGMENT) (NMR, 15 STRUCTURES)	
916	1urk		243	369	1.1e-46	0.46	1.00		PLASMINOGEN ACTIVATION PLASMINOGEN ACTIVATOR (UROKINASE-TYPE) (AMINO TERMINAL FRAGMENT) (NMR, 15 STRUCTURES)	
916	1xka	L	10	105	8.4e-16			50.64	BLOOD COAGULATION FACTOR XA; CHAIN: L, C; C <sub>3</sub>	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
916	2fm2		97	148	1.7e-20	0.72	1.00		FIBRONECTIN; CHAIN:	GLYCOPROTEIN GLYCOPROTEIN,

523

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
916	2hpp	P	11	92	1.4e-17			60.55	HYDROLASE(SERINE PROTEINASE) ALPHA- THROMBIN (E.C.3.4.21.5) COMPLEX WITH 2HPP 3 D-PHE-PRO-ARG- CHLOROMETHYLKETO NE (PPACK) CHLOROMETHYLKETO NE 2HPP 4 REPLACED BY A METHYLENE GROUP AND BOVINE PROTHROMBIN 2HPP 5 FRAGMENT 2 2HPP 6	FIBRONECTIN, TYPE TWO MODULE, NMR STRUCTURE, 2 GLYCOSYLATED PROTEIN, COLLAGEN
916	2hpq	P	11	92	4.2e-17			52.00	HYDROLASE(SERINE PROTEINASE) ALPHA- THROMBIN (E.C.3.4.21.5) COMPLEX WITH 2HPQ 3 D-PHE-PRO-ARG- CHLOROMETHYLKETO NE (PPACK) CHLOROMETHYLKETO NE 2HPQ 4 REPLACED BY A METHYLENE GROUP AND HUMAN PROTHROMBIN 2HPQ 5 FRAGMENT 2 2HPQ 6	
916	2pf1		148	268	1.1e-20	0.32	-0.18		HYDROLASE(SERINE PROTEINASE) PROTHROMBIN FRAGMENT 1 (RESIDUES 1 - 156) 2PF1	

524

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
916	2pf1		1	102	2.8e-24			50.59	3 HYDROLASE(SERINE PROTEINASE) PROTHROMBIN FRAGMENT 1 (RESIDUES 1 - 156) 2PF1	
916	2pf1		265	368	1.7e-31	0.77	0.83		3 HYDROLASE(SERINE PROTEINASE) PROTHROMBIN FRAGMENT 1 (RESIDUES 1 - 156) 2PF1	
916	2pf2		148	263	8.4e-20	0.15	-0.20		3 HYDROLASE(SERINE PROTEINASE) PROTHROMBIN FRAGMENT 1 (RESIDUES 1 - 156) COMPLEX WITH 2PF2 3 CALCIUM 2PF2 4	
916	2sta	E	11	247	2.8e-95			184.42	TRYPSIN; CHAIN: E; TRYPSIN INHIBITOR; CHAIN: 1	HYDROLASE/HYDROLASE INHIBITOR SERINE PROTEINASE, TRYPSIN INHIBITOR
916	2sta	E	445	681	1.4e-95			187.74	TRYPSIN; CHAIN: E; TRYPSIN INHIBITOR; CHAIN: 1	HYDROLASE/HYDROLASE INHIBITOR SERINE PROTEINASE, TRYPSIN INHIBITOR
916	2sta	E	445	682	1.4e-95	0.96	1.00		TRYPSIN; CHAIN: E; TRYPSIN INHIBITOR; CHAIN: 1	HYDROLASE/HYDROLASE INHIBITOR SERINE PROTEINASE, TRYPSIN INHIBITOR
916	3ktiv		10	92	4.2e-32			75.13	APOLIPROTEIN; CHAIN: NULL;	KRINGLE KRINGLE, LYSINE BINDING SITE, APOLIPROTEIN(A)
916	3ktiv		285	367	4.2e-29	0.34	1.00		APOLIPROTEIN; CHAIN: NULL;	KRINGLE KRINGLE, LYSINE BINDING SITE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
916	3kiv		77	153	2.8e-21	0.16	-0.09		APOLIPOROTEIN; CHAIN: NULL;	APOLIPOROTEIN(A) KRINGLE KRINGLE, LYSINE BINDING SITE,
916	5hpq	A	10	95	1.3e-25			67.77	PLASMINOGEN; CHAIN: A, B;	SERINE PROTEASE SERINE PROTEASE, KRINGLE 5, HUMAN PLASMINOGEN, FIBRINOLYSIS
916	5pfp		11	248	4.2e-92			185.89	BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
916	5pfp		445	682	7e-93	0.89	1.00		BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
916	5pfp		445	682	7e-93			188.52	BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
917	1d4v	A	903	1026	3e-14	0.10	-1202.08		TNF-RELATED APOPTOSIS INDUCING LIGAND; CHAIN: B; DEATH RECEPTOR 5; CHAIN: A;	APOPTOSIS TRAIL; DR5; LIGAND-RECEPTOR COMPLEX, TRIMERIC JELLY-ROLL, TNF-R 2 SUPERFAMILY, APOPTOSIS
917	1d4v	A	959	1075	3e-14	0.12	-1202.08		TNF-RELATED APOPTOSIS INDUCING LIGAND; CHAIN: B; DEATH RECEPTOR 5; CHAIN: A;	APOPTOSIS TRAIL; DR5; LIGAND-RECEPTOR COMPLEX, TRIMERIC JELLY-ROLL, TNF-R 2 SUPERFAMILY, APOPTOSIS
917	1emm		452	510	2.8e-09	0.19	-1202.08		FIBRILIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILIN-1 FRAGMENT, MATRIX PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
917	1ext	A	280	475	1.5e-19	0.04	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
917	1ext	A	819	990	4.5e-24	0.24	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
917	1ext	A	867	1054	7.5e-25	0.25	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
917	1g40	A	786	1028	6e-20	0.29	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
917	1g40	A	815	1085	1.5e-21	0.00	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
917	1g44	A	780	916	1.5e-12	0.08	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
917	1g44	B	786	1095	1.5e-19	0.14	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
917	1g44	B	925	1140	1.2e-21	-0.00	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
917	1g44	B	984	1143	4.5e-14	0.02	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
917	1g44	C	877	1140	3e-18	0.10	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
917	1igr	A	218	554	1.4e-33	0.19	-1202.08		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
917	1igr	A	781	1128	7.5e-30	0.19	-1202.08		INSULIN-LIKE GROWTH	HORMONE RECEPTOR HORMONE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
917	1klo		1051	1193	2.8e-15	0.13	-1202.08		FACTOR RECEPTOR 1; CHAIN: A;	RECEPTOR, INSULIN RECEPTOR FAMILY
917	1klo		282	470	1.5e-33	0.29	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		294	462	5.6e-24	0.15	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		349	520	1.5e-32	0.37	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		365	519	2.8e-26	0.35	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		412	561	4.5e-29	0.44	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		424	569	1.3e-24	0.43	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		472	624	1.4e-17	0.23	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		786	925	1.4e-24	0.75	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		834	985	4.5e-28	0.61	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		839	986	2.8e-26	0.60	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		924	1099	3e-38	0.68	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		928	1091	2.8e-30	0.28	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1klo		987	1139	2.8e-26	0.57	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
917	1ncf	A	819	967	3e-14	0.11	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; INCF 4 CHAIN: A, B; INCF 5	SIGNALING PROTEIN TYPE I RECEPTOR, STNFR1; INCF 8 BINDING PROTEIN, CYTOKINE INCF 19
917	1ncf	A	911	1071	1.5e-17	0.19	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; INCF 4 CHAIN: A, B; INCF 5	SIGNALING PROTEIN TYPE I RECEPTOR, STNFR1; INCF 8 BINDING PROTEIN, CYTOKINE INCF 19
917	1pp2	L	413	558	1.5e-09	0.01	-1202.08		HYDROLASE CALCIUM-FREE PHOSPHOLIPASE A=2= (E.C.3.1.1.4) 1PP2 4	
917	1qub	A	414	569	7.5e-12	0.36	-1202.08		HUMAN BETA2-GLYCOPROTEIN1; CHAIN: A;	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2-N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
917	1qub	A	802	1116	7.5e-29	0.25	-1202.08		HUMAN BETA2-	MEMBRANE ADHESION SHORT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GLYCOPROTEIN I; CHAIN: A;	CONSENSUS REPEAT, SUSP. COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
917	1skz		1036	1142	1.5e-14	0.06	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
917	1skz		346	476	3e-15	0.08	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
917	1skz		409	526	3e-16	-0.00	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
917	1skz		786	882	1.5e-17	0.13	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
917	1skz		830	931	1.1e-15	0.75	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS



Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
917	1skz		877	989	6e-14	0.26	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
917	1skz		984	1097	3e-11	0.09	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
917	1tle		282	349	9e-11	0.21	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN LAMININ-TYPE EGF-LIKE; GLYCOPROTEIN, EXTRACELLULAR MATRIX PROTEIN, NIDOGEN BINDING, 2 LE-MODULE
917	1tle		411	472	1.2e-09	0.17	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN LAMININ-TYPE EGF-LIKE; GLYCOPROTEIN, EXTRACELLULAR MATRIX PROTEIN, NIDOGEN BINDING, 2 LE-MODULE
917	9wga	A	1002	1195	4.2e-15	0.13	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
917	9wga	A	311	526	1.5e-16	0.12	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
917	9wga	A	370	554	7e-13	0.28	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
917	9wga	A	376	556	1.5e-23	0.30	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
917	9wga	A	425	608	2.8e-14	0.21	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
917	9wga	A	453	650	1.4e-13	0.12	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
917	9wga	A	793	978	5.6e-12	0.44	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
917	9wga	A	798	989	3e-20	0.22	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
917	9wga	A	863	1081	1.5e-19	0.06	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
918	lapq		27	69	1e-10	0.27	-1202.08		COMPLEMENT PROTEASE CIR; CHAIN: NULL;	COMPLEMENT, EGF, CALCIUM BINDING, SERINE PROTEASE
918	laut	L	29	124	1.1e-21	0.28	-1202.08		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
918	1aut	L	323	415	1.5e-19	0.12	-1202.08		ACTIVATED PROTEIN C <sub>2</sub> ; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COAGULATION/INHIBITOR)
										COMPLEX (BLOOD COAGULATION/INHIBITOR)
										COMPLEX (BLOOD COAGULATION/INHIBITOR)
918	1bk9		284	407	1.1e-25	0.36	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE HYDROLASE, PHOSPHOLIPASE A2, PLATELET AGGREGATION INHIBITOR, 2 PPPB HEADER MODRES
918	1dan	L	217	330	7.5e-28	0.05	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DEFRCMK) WITH CHAIN: C <sub>2</sub>	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
918	1dan	L	29	125	3e-23	0.58	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DEFRCMK) WITH CHAIN: C <sub>2</sub>	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
918	1dan	L	364	460	1.4e-09	0.36	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
918	1dan	L	7	119	1.3e-14	0.00	-1202.08		NE (DFFRCKM) WITH CHAIN: C;	
									BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCKM) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
918	1dqb	A	67	151	4.5e-24	0.03	-1202.08		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
918	1dva	L	28	123	7.5e-23	0.22	-1202.08		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
918	1dva	L	364	460	1.4e-09	0.50	-1202.08		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
918	1dx5	I	112	243	9e-19	0.38	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H; THROMBIN LIGHT	PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
918	1dx5	I	165	282	3e-19	0.75	-1202.08		CHAIN: CHAIN: A, B, C, D: THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H; THROMBIN LIGHT	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
918	1dx5	I	283	404	9.8e-20	0.17	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D: THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H; THROMBIN LIGHT	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
918	1dx5	I	28	147	4.2e-15	0.06	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D: THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H; THROMBIN LIGHT	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
918	1emn		28	88	3e-17	0.38	-1202.08		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILIN-1 FRAGMENT, MATRIX PROTEIN
918	1esl		28	142	3e-14	0.03	-1202.08		CELL ADHESION PROTEIN E-SELECTIN (LECTIN AND EGF DOMAINS, RESIDUES 1-157) 1ESL.3 (FORMERLY KNOWN AS ELAM-1) 1ESL.4	
918	1ext	A	24	92	1.5e-11	0.16	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
918	1ext	A	33	189	1e-30	0.08	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
918	1ext	A	57	219	1.5e-27	0.50	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
918	1fte	A	29	74	1.5e-12	0.31	-1202.08		BLOOD COAGULATION FACTOR VII; CHAIN: A;	BLOOD CLOTTING FACTOR VII, BLOOD COAGULATION, EGF-LIKE DOMAIN, BLOOD 2 CLOTTING
918	1fak	L	28	125	1.5e-23	0.51	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
918	1fak	L	321	410	1.1e-20	0.14	-1202.08		BLOOD COAGULATION	BLOOD CLOTTING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, BGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
918	1fak	L	364	460	1.4e-09	0.32	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, BGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
918	1g40	A	151	393	3e-21	0.28	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
918	1g40	A	68	275	1.5e-23	0.02	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
918	1g44	A	119	314	7.5e-18	0.28	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
918	1g44	B	46	282	1.5e-28	0.02	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
918	1g44	C	110	314	4.5e-18	0.12	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
918	1hj7	A	161	243	7.5e-17	0.17	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
918	1hj7	A	286	356	1.5e-26	0.12	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
918	1hj7	A	31	110	6e-24	0.31	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
918	1hj7	A	327	404	6e-22	0.30	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
918	1hj7	A	72	151	1.5e-25	0.36	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
918	1igr	A	56	400	7.5e-37	0.10	-1202.08		INSULIN-LIKE GROWTH FACTOR RECEPTOR I; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
918	1kig	L	247	295	4.5e-17	0.02	-1202.08		FACTOR XA; CHAIN: H, I; ANTI COAGULANT PEPTIDE; CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) RTAP; GLYCOPROTEIN, SERINE PROTEASE, PLASMA, BLOOD COAGULATION, 2 COMPLEX (PROTEASE/INHIBITOR)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
918	1klo		507	650	2.8e-09	0.04	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
918	1pfx	L	210	348	1.5e-21	0.06	-1202.08		FACTOR IXA; CHAIN: C; L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
918	1pfx	L	364	460	1.4e-09	0.45	-1202.08		FACTOR IXA; CHAIN: C; L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
918	1pfx	L	44	185	4.5e-25	0.02	-1202.08		FACTOR IXA; CHAIN: C; L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
918	1pfx	L	6	119	2.8e-11	0.05	-1202.08		FACTOR IXA; CHAIN: C; L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
918	1qfk	L	368	460	2.8e-08	0.54	-1202.08		COAGULATION FACTOR VIIA (LIGHT CHAIN);	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: L; COAGULATION FACTOR VIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	PROTEASE
918	1qup	A	240	480	7.5e-20	0.11	-1202.08		HUMAN BETA2- GLYCOPROTEIN I; CHAIN: A;	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
918	1sfp		704	817	3e-21	0.72	-1202.08		ASFP; CHAIN: NULL;	SPERMADHESIN ACIDIC SEMINAL PROTEIN; SPERMADHESIN, BOVINE SEMINAL PLASMA PROTEIN, ACIDIC 2 SEMINAL FLUID PROTEIN, ASFP, CUB DOMAIN, X- RAY CRYSTAL 3 STRUCTURE, GROWTH FACTOR
918	1sfp		705	822	1.4e-08	0.84	-1202.08		ASFP; CHAIN: NULL;	SPERMADHESIN ACIDIC SEMINAL PROTEIN; SPERMADHESIN, BOVINE SEMINAL PLASMA PROTEIN, ACIDIC 2 SEMINAL FLUID PROTEIN, ASFP, CUB DOMAIN, X- RAY CRYSTAL 3 STRUCTURE, GROWTH FACTOR
918	1spp	A	708	817	6e-21	0.16	-1202.08		MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	COMPLEX (SEMINAL PLASMA PROTEIN/SP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN/SP)
918	1spp	B	708	821	1.2e-23	0.17	-1202.08		MAJOR SEMINAL	COMPLEX (SEMINAL PLASMA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	PROTEIN(SPP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN(SPP))
918	1vap	A	285	407	1e-24	0.13	-1202.08		PHOSPHOLIPASE A2; CHAIN: A, B;	LIPID DEGRADATION PHOSPHOLIPASE A2, LIPID DEGRADATION, HYDROLASE
918	9wga	A	361	540	1.4e-13	0.17	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
918	9wga	A	393	552	2.8e-14	0.19	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
918	9wga	A	435	591	1.4e-15	0.25	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
920	1ae7		239	367	9e-18	0.12	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE PLA2, PHOSPHATIDE SN-2 ACYLHYDROLASE; HYDROLASE, PHOSPHOLIPASE A2, LIPID DEGRADATION, PRESTYNAPTIC 2 NEUROTOXIN, VENOM
920	1ae7		275	395	7.5e-19	0.12	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE PLA2, PHOSPHATIDE SN-2 ACYLHYDROLASE; HYDROLASE, PHOSPHOLIPASE A2, LIPID DEGRADATION, PRESTYNAPTIC 2 NEUROTOXIN, VENOM
920	1aut	L	10	123	8.4e-10			54.46	ACTIVATED PROTEIN C;	COMPLEX (BLOOD

540

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
920	1aut	L	1	96	5.6e-17			58.42	ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
920	1aut	L	223	333	1.5e-16	0.57	-1202.08		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
920	1aut	L	252	367	1e-14	0.07	-1202.08		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
920	1aut	L	295	409	6e-13	0.45	-1202.08		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1aut	L	79	165	1.2e-22	0.54	-1202.08		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO- MAL; CHAIN: P;	PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
920	1aut	L	8	108	1.4e-21			65.70	ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO- MAL; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
920	1aut	L	8	95	4.5e-21	0.13	-1202.08		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO- MAL; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
920	1bk9		124	270	6e-09	0.08	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE HYDROLASE, PHOSPHOLIPASE A2, PLATELET AGGREGATION INHIBITOR, 2 PBBP HEADER MODRES
920	1bk9		239	354	9e-18	0.01	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE HYDROLASE, PHOSPHOLIPASE A2, PLATELET AGGREGATION INHIBITOR, 2

542

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1bk9		44	154	6e-21	0.21	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	PBPB HEADER MODRES HYDROLASE HYDROLASE, PHOSPHOLIPASE A2, PLATELET AGGREGATION INHIBITOR, 2 PBPB HEADER MODRES
920	1c1g	A	185	452	1.3e-11			67.87	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
920	1c1g	A	24	304	4.2e-13			65.53	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
920	1c2a	A	10	147	2.8e-06			61.18	BOWMAN-BIRK TRYPSIN INHIBITOR; CHAIN: A	HYDROLASE INHIBITOR ALL- BETA STRUCTURE, HYDROLASE INHIBITOR
920	1c2a	A	280	430	1.5e-11	0.51	-1202.08		BOWMAN-BIRK TRYPSIN INHIBITOR; CHAIN: A	HYDROLASE INHIBITOR ALL- BETA STRUCTURE, HYDROLASE INHIBITOR
920	1dan	L	1	110	1.3e-22			54.26	BLOOD COAGULATION FACTOR VIIA; CHAIN: L; H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETO NE (DIFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
920	1dan	L	252	411	7.5e-15	0.12	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETO NE (DIFFRCMK) WITH	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
920	1dan	L	363	448	4.2e-12	0.41	-1202.08		CHAIN: C; BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DEFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
920	1dan	L	9	135	1.4e-31	0.22	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DEFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
920	1dqb	A	82	168	3e-22	0.04	-1202.08		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
920	1dqb	A	8	86	1.5e-16	0.62	-1202.08		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
920	1dva	L	1	95	2.8e-16			51.15	DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
920	1dva	L	276	366	1.5e-12	0.23	-1202.08		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN:	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE

544

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									H, I, DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	COMPLEX
920	1dva	L	363	448	4.2e-12	0.34	-1202.08		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
920	1dva	L	3	99	1.3e-13			52.84	DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
920	1dva	L	81	168	1.5e-22	0.28	-1202.08		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
920	1dva	L	8	98	4.5e-18	0.23	-1202.08		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	
920	1dva	L	9	110	1.3e-22			69.15	DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
920	1dx5	I	1	103	7e-18			54.30	THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
920	1dx5	I	1	129	1.1e-12			52.74	THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
920	1dx5	I	231	353	6e-16	0.18	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
920	1dx5	I	41	163	7.5e-25	0.43	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
920	1dx5	I	8	112	4.2e-15	0.50	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
920	1dx5	I	8	122	1.4e-26	0.69	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
920	1emm		10	93	8.4e-20			71.76	FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY,

547

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1enn		1	82	2.8e-17			52.08	FIBRILLIN; CHAIN: NULL;	DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
										MATRIX PROTEIN
920	1enn		358	434	9.8e-16	0.30	-1202.08		FIBRILLIN; CHAIN: NULL;	EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
920	1enn									MATRIX PROTEIN
										EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
920	1esl		252	395	9e-16	0.19	-1202.08		CELL ADHESION PROTEIN E-SELECTIN (LECTIN AND EGF DOMAINS, RESIDUES 1-157) IESL 3 (FORMERLY KNOWN AS ELAM-1)	
									IESL 4	
920	1esl		280	439	7.5e-15	0.33	-1202.08		CELL ADHESION PROTEIN E-SELECTIN (LECTIN AND EGF DOMAINS, RESIDUES 1-157) IESL 3 (FORMERLY KNOWN AS ELAM-1)	
									IESL 4	
920	1ext	A	223	389	6e-26	0.50	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR;	SIGNALING PROTEIN BINDING PROTEIN, CYTOKINE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1ext	A	23	168	3e-29	0.55	-1202.08		CHAIN: A, B; TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
920	1ext	A	282	451	3e-19	0.42	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
920	1ext	A	68	258	3e-22	0.09	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
920	1ext	A	8	141	1.5e-25	0.17	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
920	1fak	L	223	331	1.5e-13	0.39	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
920	1fak	L	279	367	3e-12	0.27	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
920	1fak	L	363	448	4.2e-12	0.26	-1202.08		BLOOD COAGULATION	BLOOD CLOTTING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
920	1fak	L	76	168	3e-24	0.20	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
920	1fak	L	8	100	9e-20	0.20	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
920	1fak	L	9	111	1.3e-22			63.80	BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO- FACTOR, RECEPTOR ENZYME, 3

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1g40	A	228	402	4.5e-19	0.21	-1202.08		5L15; CHAIN: I;	INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
920	1g40	A	232	439	1.3e-25	0.12	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
920	1g40	A	45	315	4.5e-19	0.09	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
920	1g40	A	58	311	4.5e-19			91.21	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
920	1g44	A	223	435	7.5e-24	0.02	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1g44	A	67	325	1.3e-20	0.15	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1g44	A	92	339	4.5e-17	0.01	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1g44	B	236	452	1.5e-23	0.36	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1g44	B	54	325	3e-22	0.18	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1g44	B	58	311	3e-22			85.82	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
920	1g44	C	276	462	7.5e-19	0.02	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1g44	C	50	314	4.5e-14	0.03	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1g44	C	82	324	6e-14	0.06	-1202.08		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
920	1hj7	A	13	93	2.8e-21			79.27	LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
920	1hj7	A	1	65	7e-21			59.44	LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
920	1hj7	A	1	66	4.2e-19			56.38	LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
920	1hj7	A	1	82	2.8e-20			60.67	LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
920	1hj7	A	44	123	3e-26			92.35	LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1hj7	A	45	123	3e-26	0.82	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
920	1hj7	A	6	82	9e-20	0.60	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
920	1hj7	A	86	163	3e-25	0.75	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
920	1igr	A	68	447	3e-32	0.02	-1202.08		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
920	1klo		15	148	1.1e-24			89.71	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
920	1klo		239	367	4.5e-27	0.50	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
920	1klo		280	413	4.5e-27			89.39	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
920	1klo		280	441	4.5e-22	0.49	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
920	1klo		323	455	1.4e-14	0.33	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
920	1klo		366	483	7e-14	0.22	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
920	1klo		77	245	5.6e-11			62.00	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
920	1nub	A	279	461	3e-13	0.00	-1202.08		BASEMENT MEMBRANE PROTEIN BM-40; CHAIN: A, B;	EXTRACELLULAR MODULE OSTEOONECTIN, SPARC, SECRETED PROTEIN ACIDIC AND EXTRACELLULAR MODULE, GLYCOPROTEIN, ANTI-ADHESIVE PROTEIN, 2 COLLAGEN BINDING, SITE-DIRECTED MUTAGENESIS, GLYCOSYLATED 3 PROTEIN MODRES



553

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMR score	SEQFOL D score	Compound	PDB annotation
920	1pfx	L	18	154	2.8e-12			56.67	FACTOR IXA; CHAIN: C <sub>1</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
920	1pfx	L	219	340	3e-17	0.16	-1202.08		FACTOR IXA; CHAIN: C <sub>1</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
920	1pfx	L	2	116	4.2e-20			58.15	FACTOR IXA; CHAIN: C <sub>1</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
920	1pfx	L	50	165	1.5e-26	0.54	-1202.08		FACTOR IXA; CHAIN: C <sub>1</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
920	1pp2	L	239	355	1.5e-17	-0.00	-1202.08		HYDROLASE CALCIUM-FREE PHOSPHOLIPASE A=2= (E.C.3.1.1.4) 1PP2 4	

554

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1qfk	L	16	112	4.2e-21			63.17	COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
920	1qfk	L	1	96	2.8e-16			54.95	COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
920	1qfk	L	363	448	4.2e-12	0.09	-1202.08		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
920	1qub	A	33	431	1.5e-26	0.06	-1202.08		HUMAN BETA2-GLYCOPROTEIN I; CHAIN: A;	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
920	1skz		11	115	1.3e-06			57.67	ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
920	1skz		239	342	3e-17	0.00	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR;

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	1skz		279	414	1.5e-20	0.20	-1202.08		ANTISTASIN; CHAIN: NULL;	ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
920	1skz		323	451	7.5e-15	0.21	-1202.08		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
920	1skz		40	150	4.2e-08			59.01	ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
920	1tpg		223	307	1.2e-14	0.08	-1202.08		T-PLASMINOGEN ACTIVATOR F1-G; 1TPG 7 CHAIN: NULL; 1TPG 8	PLASMINOGEN ACTIVATION
920	1vap	A	239	355	4.5e-17	0.15	-1202.08		PHOSPHOLIPASE A2; CHAIN: A, B;	LIPID DEGRADATION PHOSPHOLIPASE A2, LIPID DEGRADATION, HYDROLASE
920	1vap	A	324	447	3e-13	0.27	-1202.08		PHOSPHOLIPASE A2; CHAIN: A, B;	LIPID DEGRADATION PHOSPHOLIPASE A2, LIPID DEGRADATION, HYDROLASE
920	1xka	L	16	110	9.8e-19			67.01	BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL ID score	Compound	PDB annotation
										PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
920	1xka	L	1	100	1.4e-17			63.53	BLOOD COAGULATION FACTOR XA; CHAIN: L, C <sub>3</sub>	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
920	1xka	L	276	366	6e-15	0.24	-1202.08		BLOOD COAGULATION FACTOR XA; CHAIN: L, C <sub>3</sub>	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
920	1xka	L	7	98	1.5e-18	0.15	-1202.08		BLOOD COAGULATION FACTOR XA; CHAIN: L, C <sub>3</sub>	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
920	1xka	L	88	165	3e-20	1.13	-1202.08		BLOOD COAGULATION FACTOR XA; CHAIN: L, C <sub>3</sub>	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
920	9wga	A	1	150	1.4e-15			73.10	LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
920	9wga	A	223	354	1.2e-20	0.48	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
920	9wga	A	232	414	1.5e-24	0.47	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
920	9wga	A	266	433	1.4e-12	0.38	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	
920	9wga	A	273	448	6e-24	0.49	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	
920	9wga	A	310	466	7e-14	0.29	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	
920	9wga	A	47	237	9.8e-15			54.16	LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	
922	1aye		1	301	0			301.57	PROCARBOXYPEPTIDASE E A2; CHAIN: NULL;	SERINE PROTEASE PCPA2; SERINE PROTEASE, ZYMOGEN, HYDROLASE
922	1aye		1	446	0			424.65	PROCARBOXYPEPTIDASE E A2; CHAIN: NULL;	SERINE PROTEASE PCPA2; SERINE PROTEASE, ZYMOGEN, HYDROLASE
922	1aye		3	446	0	0.54	-1202.08		PROCARBOXYPEPTIDASE E A2; CHAIN: NULL;	SERINE PROTEASE PCPA2; SERINE PROTEASE, ZYMOGEN, HYDROLASE
922	1dtd	A	144	446	0	0.85	-1202.08		CARBOXYPEPTIDASE A2; CHAIN: A; METALLOCARBOXYPEP TIDASE INHIBITOR; CHAIN: B	HYDROLASE/HYDROLASE INHIBITOR CARBOXYPEPTIDASE A2, LEECH CARBOXYPEPTIDASE INHIBITOR
922	1dtd	A	144	446	0			364.24	CARBOXYPEPTIDASE A2; CHAIN: A; METALLOCARBOXYPEP TIDASE INHIBITOR;	HYDROLASE/HYDROLASE INHIBITOR CARBOXYPEPTIDASE A2, LEECH CARBOXYPEPTIDASE INHIBITOR

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
922	1dtd	A	6	301	0			346.47	CHAIN: B CARBOXYPEPTIDASE A2; CHAIN: A; METALLOCARBOXYPEPTIDASE INHIBITOR; CHAIN: B	HYDROLASE/HYDROLASE INHIBITOR CARBOXYPEPTIDASE A2, LEECH CARBOXYPEPTIDASE INHIBITOR
922	1pca		1	301	0			312.30	HYDROLASE(C- TERMINAL PEPTIDASE) PROCARBOXYPEPTIDASE A (E.C.3.4.12.2) 1PCA 3	
922	1pca		1	446	0			444.93	HYDROLASE(C- TERMINAL PEPTIDASE) PROCARBOXYPEPTIDASE A (E.C.3.4.12.2) 1PCA 3	
922	2ctc		138	444	0			388.28	HYDROLASE(C- TERMINAL PEPTIDASE) CARBOXYPEPTIDASE A (E.C.3.4.17.1) COMPLEX WITH L-PHENYL 2CTC 3 LACTATE (L-O-PHE) 2CTC 4	
922	2ctc		139	444	0	0.92	-1202.08		HYDROLASE(C- TERMINAL PEPTIDASE) CARBOXYPEPTIDASE A (E.C.3.4.17.1) COMPLEX WITH L-PHENYL 2CTC 3 LACTATE (L-O-PHE) 2CTC 4	
922	2ctc		1	301	0			376.85	HYDROLASE(C- TERMINAL PEPTIDASE) CARBOXYPEPTIDASE A (E.C.3.4.17.1) COMPLEX WITH L-PHENYL 2CTC 3 LACTATE (L-O-PHE) 2CTC 4	

559

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
927	1bkf		1	92	9.8e-32			61.12	FK506 BINDING PROTEIN; CHAIN: NULL;	ISOMERASE FKBP; ISOMERASE, ROTAMASE
927	1bkf		30	125	1.4e-31			68.15	FK506 BINDING PROTEIN; CHAIN: NULL;	ISOMERASE FKBP; ISOMERASE, ROTAMASE
927	1bkf		34	125	1.4e-31	0.49	0.96		FK506 BINDING PROTEIN; CHAIN: NULL;	ISOMERASE FKBP; ISOMERASE, ROTAMASE
927	1c9h	A	30	125	1.4e-33			53.32	FKBP12.6; CHAIN: A;	IMMUNE SYSTEM CALCINEURIN; FKBP12, RAPAMYCIN, COMPLEX, RYANODINE RECEPTOR
927	1c9h	A	34	125	1.4e-33	0.47	0.92		FKBP12.6; CHAIN: A;	IMMUNE SYSTEM CALCINEURIN; FKBP12, RAPAMYCIN, COMPLEX, RYANODINE RECEPTOR
927	1c9h	A	34	125	2.8e-28	0.50	0.76		FKBP12.6; CHAIN: A;	IMMUNE SYSTEM CALCINEURIN; FKBP12, RAPAMYCIN, COMPLEX, RYANODINE RECEPTOR
927	1fd9	A	16	126	4.2e-35	0.21	0.15		MACROPHAGE INFECTIVITY POTENTIATOR PROTEIN; CHAIN: A;	ISOMERASE MIP, PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, PIASE, FKBP DOMAIN, LONG ALPHA HELIX, DIMERISATION VIA HELICAL 2 INTERACTIONS
927	1fd9	A	26	126	8.5e-32	0.52	0.64		MACROPHAGE INFECTIVITY POTENTIATOR PROTEIN; CHAIN: A;	ISOMERASE MIP, PEPTIDYL-PROLYL CIS-TRANS ISOMERASE, PIASE, FKBP DOMAIN, LONG ALPHA HELIX, DIMERISATION VIA HELICAL 2 INTERACTIONS
927	1pbk		1	92	2.8e-26			59.85	FKBP25; CHAIN: NULL;	ISOMERASE FKBP25 C-TERMINAL DOMAIN; FKBP12 HOMOLOGOUS DOMAIN OF HFKBP25, ISOMERASE
927	1pbk		29	125	1e-31			65.99	FKBP25; CHAIN: NULL;	ISOMERASE FKBP25 C-TERMINAL DOMAIN; FKBP12 HOMOLOGOUS DOMAIN OF HFKBP25, ISOMERASE
927	1pbk		30	125	1e-31	0.73	0.99		FKBP25; CHAIN: NULL;	ISOMERASE FKBP25 C-TERMINAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
927	1rot		24	125	5.1e-33	0.50	0.98		FKBP59-I; CHAIN: NULL;	DOMAIN: FKBP12 HOMOLOGOUS DOMAIN OF HFKBP25, ISOMERASE
927	1rot		24	126	5.1e-33			60.50	FKBP59-I; CHAIN: NULL;	ROTAMASE (ISOMERASE) FKBP52 OR HSP56; ROTAMASE (ISOMERASE), DOMAIN I (N-TERM) OF A 59 KDA, 2 FK506-BINDING PROTEIN, PEPTIDYL PROLYL CIS-TRANS ISOMERASE
927	1rot		26	125	1.4e-32	0.53	0.92		FKBP59-I; CHAIN: NULL;	ROTAMASE (ISOMERASE) FKBP52 OR HSP56; ROTAMASE (ISOMERASE), DOMAIN I (N-TERM) OF A 59 KDA, 2 FK506-BINDING PROTEIN, PEPTIDYL PROLYL CIS-TRANS ISOMERASE
927	1yat		1	92	4.2e-30			63.65	BINDING PROTEIN FK-506 BINDING PROTEIN (12 KD, YEAST) COMPLEX WITH 1YAT 3 FK-506 1YAT 4	
927	1yat		21	124	5.1e-33	0.55	0.99		BINDING PROTEIN FK-506 BINDING PROTEIN (12 KD, YEAST) COMPLEX WITH 1YAT 3 FK-506 1YAT 4	
927	1yat		22	125	5.1e-33			72.40	BINDING PROTEIN FK-506 BINDING PROTEIN (12 KD, YEAST) COMPLEX WITH 1YAT 3	



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQOL D score	Compound	PDB annotation
927	1yat		31	124	2.8e-30	0.32	0.94		FK-506 1YAT 4 BINDING PROTEIN FK-506 BINDING PROTEIN (12 KD, YEAST) COMPLEX WITH 1YAT 3 FK-506 1YAT 4	
930	1dii	A	173	647	2.8e-55			144.75	P-CRESOL METHYLHYDROXYLAS E; CHAIN: A, B; P-CRESOL METHYLHYDROXYLAS E; CHAIN: C, D;	OXIDOREDUCTASE PCMH; PCMH; FLAVOCYTOCHROME, ELECTRON-TRANSFER, FAD, HEME
930	1e8g	A	167	647	5.6e-45			137.59	VANILLYL-ALCOHOL OXIDASE; CHAIN: A, B;	OXIDOREDUCTASE ARYL- ALCOHOL OXIDASE, 4- ALLYLPHENOL OXIDASE; FLAVOENZYME, SPECIFICITY
930	1f0x	A	182	648	5.6e-27			84.21	D-LACTATE DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE DLDH; OXIDOREDUCTASE
935	1lba		401	546	1.5e-47	0.57	0.70		HYDROLASE(ACTING ON LINEAR AMIDES) LYSOZYME (E.C.3.5.1.28) MUTANT WITH ALA 6 REPLACED BY LYS 1LBA 3 AND RESIDUES 2 -5 DELETED (DEL(2- 5),A6K) 1LBA 4	
935	1lba		429	542	2.8e-25	0.39	0.96		HYDROLASE(ACTING ON LINEAR AMIDES) LYSOZYME (E.C.3.5.1.28) MUTANT WITH ALA 6 REPLACED BY LYS 1LBA 3 AND RESIDUES 2 -5 DELETED (DEL(2- 5),A6K) 1LBA 4	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
937	14ps	B	1	231	0			251.78	14-3-3 PROTEIN ZETA/DELTA; CHAIN: A, B; PHOSHOPEPTIDE; CHAIN: Q, R;	COMPLEX (NEURONE/PEPTIDE) COMPLEX (NEURONE/PEPTIDE), 14-3-3, PHOSHOPEPTIDE, 2 SIGNAL TRANSDUCTION
937	14ps	B	43	278	0			250.78	14-3-3 PROTEIN ZETA/DELTA; CHAIN: A, B; PHOSHOPEPTIDE; CHAIN: Q, R;	COMPLEX (NEURONE/PEPTIDE) COMPLEX (NEURONE/PEPTIDE), 14-3-3, PHOSHOPEPTIDE, 2 SIGNAL TRANSDUCTION
937	14ps	B	53	278	0	0.56	1.00		14-3-3 PROTEIN ZETA/DELTA; CHAIN: A, B; PHOSHOPEPTIDE; CHAIN: Q, R;	COMPLEX (NEURONE/PEPTIDE) COMPLEX (NEURONE/PEPTIDE), 14-3-3, PHOSHOPEPTIDE, 2 SIGNAL TRANSDUCTION
937	14ps	B	58	277	6.8e-96	0.56	1.00		14-3-3 PROTEIN ZETA/DELTA; CHAIN: A, B; PHOSHOPEPTIDE; CHAIN: Q, R;	COMPLEX (NEURONE/PEPTIDE) COMPLEX (NEURONE/PEPTIDE), 14-3-3, PHOSHOPEPTIDE, 2 SIGNAL TRANSDUCTION
937	14do	A	1	227	4.2e-90			180.87	14-3-3 PROTEIN ZETA; CHAIN: A, B, C, D;	SIGNAL TRANSDUCTION SIGNAL
937	14do	A	43	274	1.4e-90			180.40	14-3-3 PROTEIN ZETA; CHAIN: A, B, C, D;	SIGNAL TRANSDUCTION SIGNAL
937	14do	A	53	274	1.4e-90	0.54	1.00		14-3-3 PROTEIN ZETA; CHAIN: A, B, C, D;	SIGNAL TRANSDUCTION SIGNAL
937	1qja	A	1	229	0			227.29	14-3-3 PROTEIN ZETA; CHAIN: A, B; PHOSHOPEPTIDE; CHAIN: Q, R	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) COMPLEX, 14-3-3, PHOSHOPEPTIDE, SIGNAL TRANSDUCTION
937	1qja	A	43	276	0			226.45	14-3-3 PROTEIN ZETA; CHAIN: A, B; PHOSHOPEPTIDE; CHAIN: Q, R	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) COMPLEX, 14-3-3, PHOSHOPEPTIDE, SIGNAL TRANSDUCTION
937	1qja	A	53	276	0	0.51	1.00		14-3-3 PROTEIN ZETA; CHAIN: A, B;	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PHOSPHOPEPTIDE; CHAIN: Q, R	COMPLEX, 14-3-3, PHOSPHOPEPTIDE, SIGNAL TRANSDUCTION
938	1a4p	A	63	150	8.5e-31	0.80	1.00		S100A10; CHAIN: A, B;	CALCIUM/PHOSPHOLIPID BINDING PROTEIN P11, CALPACTIN LIGHT CHAIN; S100 FAMILY, EF-HAND PROTEIN, LIGAND OF ANNEXIN II, 2 CALCIUM/PHOSPHOLIPID BINDING PROTEIN
938	1a4p	A	63	153	8.5e-31			69.95	S100A10; CHAIN: A, B;	CALCIUM/PHOSPHOLIPID BINDING PROTEIN P11, CALPACTIN LIGHT CHAIN; S100 FAMILY, EF-HAND PROTEIN, LIGAND OF ANNEXIN II, 2 CALCIUM/PHOSPHOLIPID BINDING PROTEIN
938	1b4c	A	1	87	2.8e-22			135.96	S-100 PROTEIN, BETA CHAIN; CHAIN: A, B;	METAL BINDING PROTEIN S100B, S100BETA; S100BETA, S100B, NMR, DIPOLAR COUPLINGS, EF-HAND, S100 2 PROTEIN, CALCIUM- BINDING PROTEIN, FOUR-HELIX BUNDLE, THREE-3 DIMENSIONAL STRUCTURE, SOLUTION STRUCTURE
938	1b4c	A	62	153	4.2e-29	0.61	1.00		S-100 PROTEIN, BETA CHAIN; CHAIN: A, B;	METAL BINDING PROTEIN S100B, S100BETA; S100BETA, S100B, NMR, DIPOLAR COUPLINGS, EF-HAND, S100 2 PROTEIN, CALCIUM- BINDING PROTEIN, FOUR-HELIX BUNDLE, THREE-3 DIMENSIONAL STRUCTURE, SOLUTION STRUCTURE
938	1b4c	A	62	153	4.2e-29			145.34	S-100 PROTEIN, BETA CHAIN; CHAIN: A, B;	METAL BINDING PROTEIN S100B, S100BETA; S100BETA, S100B,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										NMR, DIPOLAR COUPLINGS, EF-HAND, S100 2 PROTEIN, CALCIUM-BINDING PROTEIN, FOUR-HELIX BUNDLE, THREE-3 DIMENSIONAL STRUCTURE, SOLUTION STRUCTURE
938	1e8a	A	63	148	1.2e-33	0.71	1.00		S100A12; CHAIN: A, B;	S100 PROTEIN CALGRANULIN C
										S100 PROTEIN, X-RAY STRUCTURE, EF-HAND, CALCIUM BINDING
938	1e8a	A	63	149	1.2e-33			76.74	S100A12; CHAIN: A, B;	S100 PROTEIN CALGRANULIN C
										S100 PROTEIN, X-RAY STRUCTURE, EF-HAND, CALCIUM BINDING
938	1g8i	A	49	139	3.4e-07	-0.11	0.27		NEURONAL CALCIUM SENSOR 1; CHAIN: A, B;	METAL BINDING PROTEIN FREQUENIN; CALCIUM BINDING-PROTEIN, EF-HAND, CALCIUM ION
938	1iku		49	139	1.7e-06	0.41	0.23		RECOVERIN; CHAIN: NULL;	CALCIUM-BINDING PROTEIN CALCIUM-MYRISTOYL SWITCH, CALCIUM-BINDING PROTEIN
938	1mh0		1	87	2.8e-22			134.69	S-100 PROTEIN; CHAIN: NULL;	CALCIUM-BINDING CALCIUM-BINDING, ZINC, METAL-BINDING, ACETYLATION
938	1mh0		63	150	1.7e-27	0.71	1.00		S-100 PROTEIN; CHAIN: NULL;	CALCIUM-BINDING CALCIUM-BINDING, ZINC, METAL-BINDING, ACETYLATION
938	1mh0		63	150	1.7e-27			137.44	S-100 PROTEIN; CHAIN: NULL;	CALCIUM-BINDING CALCIUM-BINDING, ZINC, METAL-BINDING, ACETYLATION
938	1mh0		63	150	4.2e-27	0.71	1.00		S-100 PROTEIN; CHAIN: NULL;	CALCIUM-BINDING CALCIUM-BINDING, ZINC, METAL-BINDING, ACETYLATION
938	1mr8	A	61	153	8.5e-31			62.97	MIGRATION	METAL TRANSPORT MRP8,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INHIBITORY FACTOR-RELATED PROTEIN 8; CHAIN: A, B;	S100A8, CALGRANULIN A CALCIUM-BINDING PROTEIN, CRYSTAL STRUCTURE, MAD, MIGRATION 2 INHIBITORY FACTOR, RELATED PROTEIN 8, S100 PROTEIN
938	1mr8	A	62	153	8.5e-31	0.24	0.75		MIGRATION INHIBITORY FACTOR-RELATED PROTEIN 8; CHAIN: A, B;	METAL TRANSPORT MRP8, S100A8, CALGRANULIN A CALCIUM-BINDING PROTEIN, CRYSTAL STRUCTURE, MAD, MIGRATION 2 INHIBITORY FACTOR, RELATED PROTEIN 8, S100 PROTEIN
938	1qls	A	63	150	1e-34	0.39	0.99		S100C PROTEIN; CHAIN: A; ANNEXIN I; CHAIN: D;	COMPLEX (LIGAND/ANNEXIN) CALGIZZARIN; S100 FAMILY, EF-HAND PROTEIN, COMPLEX (LIGAND/ANNEXIN), 2 LIGAND OF ANNEXIN II, CALCIUM/PHOSPHOLIPID BINDING PROTEIN
938	1qls	A	63	153	1e-34			80.67	S100C PROTEIN; CHAIN: A; ANNEXIN I; CHAIN: D;	COMPLEX (LIGAND/ANNEXIN) CALGIZZARIN; S100 FAMILY, EF-HAND PROTEIN, COMPLEX (LIGAND/ANNEXIN), 2 LIGAND OF ANNEXIN II, CALCIUM/PHOSPHOLIPID BINDING PROTEIN
940	12e8	H	133	310	9.8e-27	0.28	0.09		2B8 (IGG1=KAPPA=) ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN IMMUNOGLOBULIN
940	12e8	H	313	504	5.6e-58	0.25	0.04		2B8 (IGG1=KAPPA=) ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN IMMUNOGLOBULIN
940	1a0q	H	25	220	1.1e-67	0.32	0.55		29G11 FAB; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY,

566

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1ad0	B	25	220	2.8e-68	0.16	0.33		FAB FRAGMENT, ANTIBODY ASB7; CHAIN: A, B, C, D;	ESTERASE
940	1ae6	H	24	219	2.8e-67	0.38	0.35		ANTIBODY CTM01; CHAIN: L, H;	IMMUNOGLOBULIN FRAGMENT
940	1afv	H	24	221	2.8e-70	0.21	-0.01		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	IMMUNOGLOBULIN, FAB FRAGMENT, HUMANISATION
940	1afv	H	312	503	2.8e-56	0.30	-0.08		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN) HIV-1 CA, HIV CA, HIV P24, P24; FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN), HIV, CAPSID PROTEIN, 2 P24
940	1ail	L	128	307	4.2e-25	0.37	0.49		FAB59.1; CHAIN: L, H; AIB142; CHAIN: P;	COMPLEX (ANTIBODY/PEPTIDE) COMPLEX (ANTIBODY/PEPTIDE), ANTIBODY, CONSTRAINED HIV-1 V3 2 LOOP PEPTIDE, IMMUNOGLOBULIN
940	1aif	A	128	311	4.2e-26	0.21	-0.02		ANTI-IDIOTYPIC FAB 409.5.3 (GG2A) FAB; CHAIN: A, B, L, H	IMMUNOGLOBULIN IMMUNOGLOBULIN, C REGION, V REGION
940	1aif	B	25	219	1.4e-66	0.08	-0.07		ANTI-IDIOTYPIC FAB 409.5.3 (GG2A) FAB; CHAIN: A, B, L, H	IMMUNOGLOBULIN IMMUNOGLOBULIN, C REGION, V REGION
940	1b2w	H	25	221	2.8e-69	0.15	0.00		ANTIBODY (LIGHT CHAIN); CHAIN: L;	IMMUNE SYSTEM IMMUNOGLOBULIN;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
940	1b2w	L	128	311	5.6e-26	0.34	-0.03		ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
940	1b6d	A	128	307	2.8e-25	0.22	0.11		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
940	1b6j	H	25	219	2.8e-67	0.19	0.16		IMMUNOGLOBULIN FAB' FRAGMENT OF MONOCLONAL ANTIBODY B72.3 1BBJ 3 (MURINE/HUMAN CHIMERA) 1BBJ 4	
940	1b1h	A	126	500	3.4e-42	0.55	0.87		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
940	1b1h	A	227	611	8.4e-25	0.27	-0.06		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
940	1b1h	A	22	396	1.4e-50			125.54	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1b1h	A	24	396	1.4e-50	0.34	0.89		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
940	1b1j	J	128	310	5.6e-26	0.10	0.09		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
940	1b1n	B	25	221	1.1e-69	-0.06	0.22		MONOCLONAL ANTIBODY MRK-16 (LIGHT CHAIN); CHAIN: A, C; MONOCLONAL ANTIBODY MRK-16 (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM IMMUNOGLOBULIN, IMMUNE SYSTEM
940	1c5c	H	24	221	2.8e-69	0.22	0.60		CHIMERIC DECARBOXYLASE ANTIBODY 21D8; CHAIN: L; CHIMERIC DECARBOXYLASE ANTIBODY 21D8; CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN, CATALYTIC ANTIBODY, CHIMERIC FAB, 2 DECARBOXYLASE, HAPTEN COMPLEX
940	1c1c	B	24	221	8.4e-70	0.50	0.41		IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
940	1c1c	B	312	503	4.2e-57	0.27	0.18		IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	
940	1cl7	I	140	222	4.2e-26	0.60	0.12		IGG1 ANTIBODY 1696 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 1696 (VARIABLE HEAVY CHAIN); CHAIN: H; IGG1 ANTIBODY 1696 (CONSTANT HEAVY CHAIN); CHAIN: J;	IMMUNE SYSTEM IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1 FAB FRAGMENT, CROSS-REACTIVITY, HIV1 PROTEASE, ENZYME 2 INHIBITION, IMMUNOGLOBULIN
940	1cl7	I	413	505	4.2e-28	0.10	-0.13		IGG1 ANTIBODY 1696 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 1696 (VARIABLE HEAVY CHAIN); CHAIN: H; IGG1 ANTIBODY 1696 (CONSTANT HEAVY CHAIN); CHAIN: J;	IMMUNE SYSTEM IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1 FAB FRAGMENT, CROSS-REACTIVITY, HIV1 PROTEASE, ENZYME 2 INHIBITION, IMMUNOGLOBULIN
940	1cf9	H	136	309	2.8e-25	0.11	0.22		FAB ANTIBODY LIGHT CHAIN; CHAIN: L; FAB ANTIBODY HEAVY CHAIN; CHAIN: H;	IMMUNE SYSTEM ANTI-PRION FAB 3F4; ANTI-PRION FAB 3F4 ANTI-PRION ANTIBODY, FAB 3F4
940	1cs6	A	122	500	1.7e-47	0.26	1.00		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
940	1cs6	A	122	504	1.7e-47			129.23	AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
940	1cs6	A	132	503	1.1e-35	0.27	0.07		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
940	1cs6	A	225	611	2.8e-35	0.21	0.27		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
940	1cs6	A	24	392	2.8e-64	0.15	0.63		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1ct8	A	128	311	4.2e-25	0.36	0.07		7C8 FAB FRAGMENT; SHORT CHAIN; CHAIN: A, C; 7C8 FAB FRAGMENT; LONG CHAIN; CHAIN: B, D	IMMUNE SYSTEM ABZYME TRANSITION STATE ANALOG, IMMUNE SYSTEM
940	1ct8	B	312	505	1.4e-56	0.18	0.31		7C8 FAB FRAGMENT; SHORT CHAIN; CHAIN: A, C; 7C8 FAB FRAGMENT; LONG CHAIN; CHAIN: B, D	IMMUNE SYSTEM ABZYME TRANSITION STATE ANALOG, IMMUNE SYSTEM
940	1dec	A	128	311	9.8e-27	0.24	0.18		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
940	1dfb	L	128	311	2.8e-26	0.03	0.01		IMMUNOGLOBULIN 3D6 FAB 1DFB 3	
940	1dgi	R	132	396	5.1e-32	0.18	0.45		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
940	1dgi	R	223	500	1.7e-23	-0.21	0.00		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
940	1dgi	R	25	308	1e-45			126.61	POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: 3; VP4; CHAIN: 4;	MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
940	1dgi	R	26	308	1e-45	-0.16	0.03		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
940	1dh2	A	144	303	9.8e-31	0.06	0.59		IMMUNOGLOBULIN LAMBDA HEAVY CHAIN; CHAIN: A; B; ENGINEERED PEPTIDE; CHAIN: E, F;	IMMUNE SYSTEM FC IGG PHAGE DISPLAY PEPTIDE
940	1dh2	A	227	393	4.2e-37	0.22	-0.01		IMMUNOGLOBULIN LAMBDA HEAVY CHAIN; CHAIN: A, B; ENGINEERED PEPTIDE; CHAIN: E, F;	IMMUNE SYSTEM FC IGG PHAGE DISPLAY PEPTIDE
940	1dh2	A	398	610	2.8e-24	0.05	-0.15		IMMUNOGLOBULIN LAMBDA HEAVY CHAIN; CHAIN: A, B; ENGINEERED PEPTIDE; CHAIN: E, F;	IMMUNE SYSTEM FC IGG PHAGE DISPLAY PEPTIDE
940	1dgg	A	128	311	8.4e-26	0.21	-0.07		ANTI-LYSOZYME ANTIBODY HYHEL-63 (LIGHT CHAIN); CHAIN: A, C; ANTI-LYSOZYME ANTIBODY HYHEL-63 (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM ANTI-LYSOZYME ANTIBODY, HYHEL-63, HEN EGG WHITE LYSOZYME
940	1e4k	A	144	303	2.8e-31	0.20	0.46		LOW AFFINITY IMMUNOGLOBULIN	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GAMMA FC RECEPTOR CHAIN: C; FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	FC, RECEPTOR, CD16, GAMMA
940	1e4k	A	221	394	1.1e-37	0.25	-0.11		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C; FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
									IGG1; CHAIN: A, B;	
940	1e4x	H	24	224	7e-73	0.22	0.58		TAB2; CHAIN: L, M; TAB2; CHAIN: H, I; CYCLIC PEPTIDE; CHAIN: P, Q	COMPLEX (ANTIBODY/ANTIGEN) CROSS-REACTIVITY, PROTEIN-PEPTIDE RECOGNITION
									TAB2; CHAIN: L, M; TAB2; CHAIN: H, I; CYCLIC PEPTIDE; CHAIN: P, Q	
940	1e4x	H	312	505	4.2e-57	0.29	0.03		TAB2; CHAIN: L, M; TAB2; CHAIN: H, I; CYCLIC PEPTIDE; CHAIN: P, Q	COMPLEX (ANTIBODY/ANTIGEN) CROSS-REACTIVITY, PROTEIN-PEPTIDE RECOGNITION
									IGG2A MONOCLONAL ANTIBODY (LIGHT CHAIN); CHAIN: L; IGG2A MONOCLONAL ANTIBODY (HEAVY CHAIN); CHAIN: H; FMDV PEPTIDE; CHAIN: P;	
940	1e4o	H	315	503	2.8e-48	0.25	0.19		IGG2A MONOCLONAL ANTIBODY (LIGHT CHAIN); CHAIN: L; IGG2A MONOCLONAL ANTIBODY (HEAVY CHAIN); CHAIN: H; FMDV PEPTIDE; CHAIN: P;	IMMUNE SYSTEM FMDV, ANTIGENIC-ANTIBODY INTERACTIONS, RGD MOTIF, G-H LOOP 2 OF VP1.
									FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
940	1e4t	C	224	395	5.1e-28	0.49	0.68		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	IMMUNE SYSTEM CATALYTIC ANTIBODY, AMIDINUM, HAPTENIC CHARGE
									CATALYTIC ANTIBODY 4B2; CHAIN: L, J; CATALYTIC ANTIBODY	
940	1F3d	H	314	505	2.8e-54	0.14	0.03		CATALYTIC ANTIBODY 4B2; CHAIN: L, J; CATALYTIC ANTIBODY	

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1f6a	A	224	399	1.2e-29	0.49	0.72		4B2: CHAIN: H, K; HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D; COMPLEX	IMMUNE SYSTEM HIGH AFFINITY IGB-FC RECEPTOR, FC(EPSILON) IGB-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGB-BINDING 2 PROTEIN, IGB ANTIBODY, IGB-FC
940	1fbi	H	312	502	4.2e-57	0.14	0.33		(ANTIBODY/ANTIGEN) FAB FRAGMENT OF THE MONOCLONAL ANTIBODY F9.13.7 (IGG1) 1FBI.3 COMPLEXED WITH LYSOZYME (E.C.3.2.1.17) 1FBI.4	
940	1fc2	D	144	303	2.8e-31	-0.02	0.46		IMMUNOGLOBULIN AND FRAGMENT B OF PROTEIN A COMPLEX 1FC2.4	
940	1fc2	D	227	394	4.2e-37	0.44	-0.06		IMMUNOGLOBULIN AND FRAGMENT B OF PROTEIN A COMPLEX 1FC2.4	
940	1fc2	D	398	612	4.2e-23	0.10	-0.19		IMMUNOGLOBULIN AND FRAGMENT B OF PROTEIN A COMPLEX 1FC2.4	
940	1fgn	H	25	219	1.3e-66	0.13	0.30		IMMUNOGLOBULIN FAB 5G9; CHAIN: L, H; IMMUNOGLOBULIN FAB 5G9; CHAIN: L, H; MONOCLONAL, MURINE,	IMMUNOGLOBULIN FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN; ANTIBODY, FAB, ANTI-1F, MONOCLONAL, MURINE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1fgn	H	314	500	5.6e-56	0.41	0.00		IMMUNOGLOBULIN FAB 5G9; CHAIN: L, H;	IMMUNOGLOBULIN
940	1fh5	H	318	503	2.8e-47	0.32	0.04		MONOCLONAL ANTIBODY MAK33; CHAIN: L; MONOCLONAL ANTIBODY MAK33; CHAIN: H;	IMMUNE SYSTEM FAB, BIP, CRYSTAL STRUCTURE
940	1fhg	L	128	311	2.8e-25	0.31	0.03		IMMUNOGLOBULIN G1 (KAPPA LIGHT CHAIN) FAB' FRAGMENT 1FIG3	
940	1fh3	A	315	501	4.2e-49	0.37	-0.02		BLUE FLUORESCENT ANTIBODY (19G2)-HEAVY CHAIN; CHAIN: H, A; BLUE FLUORESCENT ANTIBODY (19G2)-LIGHT CHAIN; CHAIN: L, B;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD
940	1fhs	L	128	311	1.4e-25	0.19	0.03		IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
940	1for	H	314	503	1.4e-53	0.53	-0.06		IMMUNOGLOBULIN IGG2A FAB FRAGMENT (FAB17-1A) (ORTHORHOMBIC	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1f5k	C	312	505	9.8e-58	0.13	0.17		CRYSTAL FORM; IFOR 3 MAJOR POLLEN ALLERGEN BET V 1-A; CHAIN: A, D, G, J; IMMUNOGLOBULIN KAPPA LIGHT CHAIN; CHAIN: B, E, H, K; ANTIBODY HEAVY CHAIN FAB; CHAIN: C, F, I, L;	IMMUNE SYSTEM BET V 1-A, BETVI ALLERGEN; BV16 FAB- FRAGMENT, KAPPA MOPC21 CODING SEQUENCE; HEAVY CHAIN OF THE MONOCLONAL ANTIBODY MST2; BET V 1, BV16 FAB FRAGMENT, ANTIBODY ALLERGEN COMPLEX
940	1fvd	A	128	311	2.8e-26	0.40	0.48		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 IFVD 3	
940	1hzh	H	139	499	2.8e-64	0.19	0.35		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
940	1hzh	H	229	612	1.4e-29	0.11	-0.03		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
940	1hzh	H	24	397	0	0.23	0.25		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
940	1hzh	H	312	611	4.2e-61	0.25	0.11		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	
940	1ilc	A	228	392	4.2e-30	0.31	-0.08		IG GAMMA-2A CHAIN C REGION; CHAIN: A, B;	IMMUNE SYSTEM IGG2A; IGG, FC
940	1ile	A	399	610	2.8e-20	0.12	-0.17		IG GAMMA-2A CHAIN C REGION; CHAIN: A, B;	IMMUNE SYSTEM IGG2A; IGG, FC
940	1igt	A	128	311	2.8e-25	0.02	0.17		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
940	1igt	B	136	499	1.1e-58	0.04	0.23		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
940	1igt	B	25	393	0	0.35	0.48		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
940	1igt	B	313	611	1.1e-67	0.08	0.03		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
940	1igy	B	136	499	1.4e-58	0.16	0.62		IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
940	1igy	B	25	392	1.4e-97	0.38	0.10		IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
940	1igy	B	313	611	5.6e-66	0.12	0.13		IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
940	1ili	A	314	503	1.3e-53	0.33	0.28		MONOCLONAL ANTIBODY G3-519 (HEAVY CHAIN); CHAIN: A; MONOCLONAL ANTIBODY G3-519 (LIGHT CHAIN); CHAIN:	IMMUNE SYSTEM FAB, BETA SHEET STRUCTURE, ANTIBODY



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1tth	B	230	500	1.4e-36	0.11	0.66		B; INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
940	1kb5	H	25	222	4.2e-70	0.23	0.23		KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T- CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
940	1kb5	L	128	311	2.8e-26	0.24	0.43		KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T- CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
940	1ma m	H	25	220	9.8e-67	-0.01	0.01		IMMUNOGLOBULIN ANTIGEN-BINDING FRAGMENT (FAB) (GG2B, KAPPA) 1MAM3	
940	1mco	H	136	499	2.8e-68	0.16	0.66		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION 1MCO 3	
940	1mco	H	227	610	1.4e-32	-0.13	0.13		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION 1MCO	

578

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
940	Imco	H	22	433	0			112.21	3 IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO	
940	Imco	H	24	393	0	0.08	0.89		3 IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO	
940	Imco	H	312	611	4.2e-57	0.31	0.11		3 IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO	
940	Imim	H	314	502	5.6e-48	0.41	0.46		3 CHIMERIC SDZ CH1621; CHAIN: H, L;	IMMUNOGLOBULIN IMMUNOGLOBULIN, C REGION
940	Inbv	H	25	220	4.2e-66	0.22	-0.12		IMMUNOGLOBULIN FAB (BV04-01) AUTOANTIBODY BINDING SINGLE- STRANDED DNA INBV 3 (UNLIGANDED FORM) INBV 4	
940	Ip1g	H	24	220	8.4e-70	0.39	0.17		IGG2A=KAPPA=, IPLG 4 CHAIN: L, H, IPLG 5	IMMUNOGLOBULIN
940	Iqfu	H	312	504	2.8e-56	0.36	0.52		HEMAGGLUTININ (HA1 CHAIN); CHAIN: A; HEMAGGLUTININ (HA2 CHAIN); CHAIN: B; IMMUNOGLOBULIN IGG1-KAPPA ANTIBODY (LIGHT CHAIN); CHAIN:	VIRAL PROTEIN/IMMUNE SYSTEM COMPLEX (HEMAGGLUTININ/IMMUNOGL OBULIN), HEMAGGLUTININ, 2 IMMUNOGLOBULIN, VIRAL PROTEIN/IMMUNE SYSTEM

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									L <sub>2</sub> IMMUNOGLOBULIN IGG1-KAPPA ANTIBODY (HEAVY CHAIN); CHAIN: H <sub>2</sub>	
940	1sm3	H	24	222	1.4e-70	0.09	0.07		SMB ANTIBODY; CHAIN: L <sub>2</sub> , H <sub>2</sub> ; PEPTIDE EPITOPE; CHAIN: P;	COMPLEX (ANTIBODY/PEPTIDE EPITOPE) ANTIBODY, PEPTIDE ANTIGEN, ANTITUMOR ANTIBODY, 2 COMPLEX (ANTIBODY/PEPTIDE EPITOPE)
940	1vge	L	128	311	2.8e-25	0.23	-0.02		TRL 9 FAB; CHAIN: L <sub>2</sub> , H <sub>2</sub>	IMMUNOGLOBULIN TRL 9, ANTI-THYROID PEROXIDASE, AUTOANTIBODY, 2 IMMUNOGLOBULIN
940	1wej	H	313	505	1.4e-56	0.32	0.05		E8 ANTIBODY; CHAIN: L <sub>2</sub> , H <sub>2</sub> ; CYTOCHROME C <sub>2</sub> ; CHAIN: F <sub>2</sub>	COMPLEX (ANTIBODY/ELECTRON TRANSPORT) FAB E8; CYT C <sub>2</sub> , ANTIGEN; IMMUNOGLOBULIN, IGG1 KAPPA, FAB FRAGMENT, HORSE 2 CYTOCHROME C <sub>2</sub> , COMPLEX (ANTIBODY/ELECTRON TRANSPORT)
940	1wej	L	128	311	1.4e-25	0.19	0.11		E8 ANTIBODY; CHAIN: L <sub>2</sub> , H <sub>2</sub> ; CYTOCHROME C <sub>2</sub> ; CHAIN: F <sub>2</sub>	COMPLEX (ANTIBODY/ELECTRON TRANSPORT) FAB E8; CYT C <sub>2</sub> , ANTIGEN; IMMUNOGLOBULIN, IGG1 KAPPA, FAB FRAGMENT, HORSE 2 CYTOCHROME C <sub>2</sub> , COMPLEX (ANTIBODY/ELECTRON TRANSPORT)
940	1wio	A	134	492	8.5e-34	0.13	-0.08		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
940	1wio	A	319	629	5.1e-18	0.22	-0.17		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	LIPOPROTEIN, POLYMORPHISM
940	1yej	H	314	502	1.3e-51	0.27	0.03		IG ANTIBODY D2.3 (LIGHT CHAIN); CHAIN: L; IG ANTIBODY D2.3 (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM ABZYME, TRANSITION STATE ANALOG, IMMUNE SYSTEM
940	25c8	H	25	219	1.4e-67	0.28	0.66		IGG 5C8; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
940	25c8	H	313	501	1.4e-56	0.26	0.07		IGG 5C8; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
940	2cgt	H	25	220	2.8e-66	-0.00	-0.07		IMMUNOGLOBULIN IGG2B (KAPPA) FAB FRAGMENT COMPLEXED WITH ANTIGEN 2CGR 3 N-(P-CYANOPHENYL)-N'-(DIPHENYLEMETHYL) GUANIDINEACETIC ACID 2CGR 4	
940	2fgw	H	25	221	1.4e-69	0.24	0.24		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 1H52' (HUH52-OZ FAB) 2FGW 4	
940	2fgw	L	128	311	7e-27	0.34	0.23		IMMUNOGLOBULIN FAB FRAGMENT OF A	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY "H52" (HUH52-OZ F4B) 2FGW 4	
941	1a4y	A	7	243	1e-26	0.41	0.96		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (R1-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
941	1a9n	A	108	242	1e-24	0.73	0.89		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	A	139	243	7e-07	0.34	1.00		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	A	25	174	1.2e-21	0.46	0.57		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	A	48	197	1.5e-25	0.69	0.96		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	A	66	173	2.8e-06	0.49	0.49		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	A	90	194	2.8e-06	0.34	0.89		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A';	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A, C; U2 B"; CHAIN: B, D;	(NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	C	108	242	6.8e-25	0.73	0.84		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	C	139	243	7e-07	0.52	1.00		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	C	30	148	1.7e-18	0.45	0.90		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1a9n	C	48	221	5.1e-26	0.67	0.95		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
941	1d0b	A	156	306	1.4e-18	0.59	0.99		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
941	1d0b	A	23	167	1.4e-23	0.55	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
941	1d0b	A	40	191	8.4e-27	0.56	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
941	1d0b	A	73	239	7e-27	1.21	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
941	1dce	A	163	267	1.4e-10	0.79	0.80		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
941	1dce	A	30	254	1.7e-35	-0.00	-0.05		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
941	1dce	A	90	196	1.4e-11	0.72	0.99		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
941	1ds9	A	23	142	8.4e-13	-0.32	0.19		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
941	1ds9	A	48	237	1.7e-18	-0.21	0.05		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
941	1ds9	A	99	238	2.8e-15	0.01	0.62		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA

584

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Pst Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
941	1fo1	A	134	196	1.4e-07	-0.09	0.96		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
941	1fo1	A	182	247	2.8e-07	0.20	0.35		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
941	1fo1	B	134	196	1.4e-07	-0.11	0.89		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
941	1fo1	B	182	247	2.8e-07	0.34	0.42		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
941	1fgv	A	44	240	1.4e-12	0.27	0.37		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
941	1fs2	A	31	240	1e-12	0.38	-0.02		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
941	1fs2	A	44	241	8.4e-11	0.23	0.27		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
941	1fs2	A	88	240	1.4e-12	0.39	0.95		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
941	1ft8	A	134	196	1.4e-07	-0.03	0.94		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH-REPEAT 2 (LRR) DOMAINS
941	1ft8	A	182	247	2.8e-07	0.76	0.36		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH-REPEAT 2 (LRR) DOMAINS
941	1yrg	A	41	242	1.7e-20	0.42	0.41		GTPASE-ACTIVATING PROTEIN RNAI_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
941	2bnh		143	335	4.2e-09	0.20	0.11		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
941	2bnh		30	242	1.7e-24	0.67	0.82		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
944	1a7c	A	170	532	2.8e-91			151.16	PLASMINOGEN ACTIVATOR INHIBITOR	COMPLEX (PROTEASE INHIBITOR/PEPTIDE) PAI-1;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									TYPE I; CHAIN: A; PENTAPEPTIDE; CHAIN: B, C	SERINE PROTEASE INHIBITOR, PAI-1, CARBOHYDRATE, INHIBITOR 2 COMPLEX, PEPTIDE
944	1a7c	A	8	370	2.8e-91			152.25	PLASMINOGEN ACTIVATOR INHIBITOR TYPE I; CHAIN: A; PENTAPEPTIDE; CHAIN: B, C	COMPLEX (PROTEASE INHIBITOR/PEPTIDE) PAI-1; SERINE PROTEASE INHIBITOR, PAI-1, CARBOHYDRATE, INHIBITOR 2 COMPLEX, PEPTIDE
944	1as4	A	15	337	0			155.82	ANTICHYMOTRYPSIN; CHAIN: A, B;	SERPIN ACT; SERPIN, SERINE PROTEASE INHIBITOR, ANTICHYMOTRYPSIN
944	1as4	A	176	489	4.2e-100	0.32	1.00		ANTICHYMOTRYPSIN; CHAIN: A, B;	SERPIN ACT; SERPIN, SERINE PROTEASE INHIBITOR, ANTICHYMOTRYPSIN
944	1as4	A	177	499	4.2e-100			154.82	ANTICHYMOTRYPSIN; CHAIN: A, B;	SERPIN ACT; SERPIN, SERINE PROTEASE INHIBITOR, ANTICHYMOTRYPSIN
944	1a7h	A	163	534	9.8e-98			168.04	HUMAN ANTIHROMBIN-III ANTIHROMBIN III 1ATH 3	
944	1a7h	A	176	532	9.8e-98	0.84	1.00		HUMAN ANTIHROMBIN-III ANTIHROMBIN III 1ATH 3	
944	1a7h	A	1	372	1.4e-97			170.36	HUMAN ANTIHROMBIN-III ANTIHROMBIN III 1ATH 3	
944	1a7h	B	168	533	2.8e-96			137.56	HUMAN ANTIHROMBIN-III ANTIHROMBIN III 1ATH 3	
944	1a7h	B	176	532	2.8e-96	0.70	1.00		HUMAN	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ANTITHROMBIN-III ANTITHROMBIN III IATH 3	
944	1ath	B	6	371	8.4e-96			138.93	HUMAN ANTITHROMBIN-III ANTITHROMBIN III IATH 3	
944	1by7	A	10	370	0			148.40	PLASMINOGEN ACTIVATOR INHIBITOR- 2; CHAIN: A;	PROTEIN BINDING PAL-2; SERPIN, PROTEIN BINDING
944	1by7	A	172	532	0			147.26	PLASMINOGEN ACTIVATOR INHIBITOR- 2; CHAIN: A;	PROTEIN BINDING PAL-2; SERPIN, PROTEIN BINDING
944	1by7	A	176	532	0	0.55	1.00		PLASMINOGEN ACTIVATOR INHIBITOR- 2; CHAIN: A;	PROTEIN BINDING PAL-2; SERPIN, PROTEIN BINDING
944	1db2	A	171	532	1.4e-91			162.23	PLASMINOGEN ACTIVATOR INHIBITOR- 1; CHAIN: A, B;	HYDROLASE INHIBITOR NATIVE SERPIN, HYDROLASE INHIBITOR
944	1db2	A	9	370	9.8e-91			163.81	PLASMINOGEN ACTIVATOR INHIBITOR- 1; CHAIN: A, B;	HYDROLASE INHIBITOR NATIVE SERPIN, HYDROLASE INHIBITOR
944	1dzg	I	136	534	2.8e-97			180.38	ANTITHROMBIN-III; CHAIN: I; ANTITHROMBIN-III; CHAIN: L;	SERPIN SERPIN
944	1dzg	I	176	532	2.8e-97	0.48	1.00		ANTITHROMBIN-III; CHAIN: I; ANTITHROMBIN-III; CHAIN: L;	SERPIN SERPIN
944	1dzg	I	1	372	4.2e-97			154.88	ANTITHROMBIN-III; CHAIN: I; ANTITHROMBIN-III; CHAIN: L;	SERPIN SERPIN

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
944	1ezx	A	12	339	2.8e-98			157.67	ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1-ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
944	1ezx	A	174	501	1.4e-99			156.36	ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1-ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
944	1ezx	A	175	501	1.4e-99	0.28	1.00		ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1-ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
944	1f0c	A	185	502	2.8e-66			104.89	ICE INHIBITOR; CHAIN: A; ICE INHIBITOR; CHAIN: B;	VIRUS/VIRAL PROTEIN CYTOKINE RESPONSE MODIFIER PROTEIN, CRMA; CYTOKINE RESPONSE MODIFIER PROTEIN, CRMA; APOPTOSIS, CASPASE INHIBITOR, PROTEASE INHIBITOR, SERPIN
944	1f0c	A	23	340	4.2e-65			106.63	ICE INHIBITOR; CHAIN: A; ICE INHIBITOR; CHAIN: B;	VIRUS/VIRAL PROTEIN CYTOKINE RESPONSE MODIFIER PROTEIN, CRMA; CYTOKINE RESPONSE MODIFIER PROTEIN, CRMA; APOPTOSIS, CASPASE INHIBITOR, PROTEASE INHIBITOR, SERPIN
944	1hle	A	13	336	4.2e-99			134.44	HYDROLASE INHIBITOR(SERINE PROTEINASE) HORSE LEUKOCYTE ELASTASE INHIBITOR (HLE) 1HLE 3	
944	1hle	A	175	489	7e-99	0.30	1.00		HYDROLASE INHIBITOR(SERINE PROTEINASE) HORSE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									LEUKOCYTE ELASTASE INHIBITOR (HLEI) 1HLE 3	
944	1hle	A	175	498	7e-99			132.76	HYDROLASE INHIBITOR(SERINE PROTEINASE) HORSE LEUKOCYTE ELASTASE INHIBITOR (HLEI) 1HLE 3	
944	1i99	I	168	532	1.5e-64			167.27	ALASERPIN; CHAIN: I; TRYPSIN II, ANIONIC; CHAIN: E; 3	HYDROLASE/HYDROLASE INHIBITOR SERPIN I; PRETRYPSINOGEN II; MICHAELIS SERPIN-PROTEASE COMPLEX INHIBITORY TRIAD
944	1i99	I	6	370	7e-81			169.22	ALASERPIN; CHAIN: I; TRYPSIN II, ANIONIC; CHAIN: E; 3	HYDROLASE/HYDROLASE INHIBITOR SERPIN I; PRETRYPSINOGEN II; MICHAELIS SERPIN-PROTEASE COMPLEX INHIBITORY TRIAD
944	1ova	A	12	370	2.8e-94			151.89	SERPIN OVALBUMIN (EGG ALBUMIN) 1OVA 3	
944	1ova	A	174	532	1.4e-96			151.52	SERPIN OVALBUMIN (EGG ALBUMIN) 1OVA 3	
944	1ova	A	176	532	1.4e-96	0.53	1.00		SERPIN OVALBUMIN (EGG ALBUMIN) 1OVA 3	
944	1qlp	A	12	372	0			171.59	ALPHA-1-ANTITRYPSIN; CHAIN: A; 3	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE, SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE

590

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQEQL D score	Compound	PDB annotation
944	1qlp	A	174	534	0			170.59	ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE; SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
944	1qlp	A	175	532	0	0.46	1.00		ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE; SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
944	1qmb	B	502	532	9.8e-09	-0.75	0.05		ALPHA-1-ANTITRYPSIN; CHAIN: A, B;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-PI; SERPIN, ANTITRYPSIN, POLYMER, CLEAVED
944	1qmn	A	14	372	0			184.90	ALPHA-1-ANTICHYMOTRYPSIN; CHAIN: A;	SERPIN AACT SERPIN, SERINE PROTEINASE INHIBITOR, PARTIAL LOOP 2 INSERTION, LOOP-SHEET POLYMERIZATION, EMPHYSEMA, DISEASE 3 MUTATION, ACUTE PHASE PROTEIN, CONFORMATIONAL DISEASE
944	1qmn	A	176	533	0	0.46	1.00		ALPHA-1-ANTICHYMOTRYPSIN; CHAIN: A;	SERPIN AACT SERPIN, SERINE PROTEINASE INHIBITOR, PARTIAL LOOP 2 INSERTION, LOOP-SHEET POLYMERIZATION,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
944	1qmm	A	176	534	0			183.22	ALPHA-1-ANTICHYMOTRYPSIN; CHAIN: A;	EMPHYSEMA, DISEASE 3 MUTATION, ACUTE PHASE PROTEIN, CONFORMATIONAL DISEASE
944	1sek		168	532	1.4e-81			143.79	SERPIN K; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, SERPIN, PROTEASE
944	1sek		6	370	2.8e-80			145.58	SERPIN K; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, SERPIN, PROTEASE
945	1t3l	A	80	188	4.2e-25	0.40	1.00		PROTEIN ARGININE METHYLTRANSFERASE PRMT3; CHAIN: A;	TRANSFERASE BETA BARREL, ROSSMANN FOLD, ARGININE METHYLTRANSFERASE
945	1t3l	A	80	213	1.7e-05	-0.30	0.23		PROTEIN ARGININE METHYLTRANSFERASE PRMT3; CHAIN: A;	TRANSFERASE BETA BARREL, ROSSMANN FOLD, ARGININE METHYLTRANSFERASE
945	1g6q	1	81	219	1.4e-25	0.03	0.57		HNRP ARGININE N- METHYLTRANSFERASE; CHAIN: 1, 2, 3, 4, 5, 6;	TRANSFERASE SAM-BINDING DOMAIN, BETA-BARREL, MIXED ALPHA-BETA, HEXAMER, 2 DIMER
947	1av1	A	1	193	2.8e-09			56.50	APOLIPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT- ACTIVATION
947	1clg	A	13	296	1.3e-29			96.22	TROPOMYOSIN; CHAIN:	CONTRACTILE PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A, B, C, D	TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
947	1c1g	A	249	532	4.2e-28			105.38	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
947	1cun	A	190	417	1.7e-13	0.09	-0.05		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
947	1cun	A	249	504	3.4e-19	0.11	0.52		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
947	1cun	A	412	536	6.8e-08	0.05	0.03		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
947	1cun	A	4	206	0.0097			56.15	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
947	1cun	A	55	271	0.00011			69.05	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
947	1dn1	B	1	190	0.0025			55.90	SYNTAXIN BINDING PROTEIN 1; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN



Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									SYNTAXIN 1A; CHAIN: B;	COMPLEX, MULTI-SUBUNIT
947	1dn1	B	366	527	1.2e-12	0.13	-0.15		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
947	1dn1	B	53	284	0.0066			63.83	SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
947	1dvp	A	475	575	3.4e-20	-0.06	0.28		HEPATOCYTE GROWTH FACTOR-REGULATED TYROSINE CHAIN: A;	TRANSFERASE HRS; HRS, VHS, FYVE, ZINC FINGER, SUPERHELIX
947	1fio	A	2	191	0.0097			62.09	SSO1 PROTEIN; CHAIN: A;	MEMBRANE PROTEIN FOUR HELIX BUNDLE, ALPHA HELIX
947	1hci	A	126	570	1.7e-30			114.13	ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN
947	1hci	A	188	553	1.7e-29	-0.13	0.22		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN
947	1hci	A	189	511	1.7e-30	-0.15	0.10		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN

594

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN
947	1hyi	A	531	580	5.1e-19	0.11	0.75		ENDOSOME-ASSOCIATED PROTEIN; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS BETA SHEET, ALPHA HELIX, ZINC CLUSTER, PTDINS(3)P
947	1hyi	A	5	68	1.1e-19			62.30	ENDOSOME-ASSOCIATED PROTEIN; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS BETA SHEET, ALPHA HELIX, ZINC CLUSTER, PTDINS(3)P
947	1quu	A	297	517	1.7e-18	0.11	-0.15		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
947	1vfy	A	531	575	6.8e-16	-0.08	0.69		PHOSPHATIDYLINOSITO L-3-PHOSPHATE BINDING FYVE CHAIN: A;	TRANSPORT PROTEIN FYVE DOMAIN, ENDOSOME MATURATION, INTRACELLULAR TRAFFICKING, 2 TRANSPORT PROTEIN
947	1zbd	B	496	603	6.8e-23	0.02	-0.02		RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-BINDING PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
950	1b6e		111	237	5.6e-26			94.09	CD94; CHAIN: NULL;	NK CELL NK CELL, RECEPTOR, C-TYPE LECTIN, C-TYPE LECTIN-LIKE, NKD
950	1b6e		1	118	7e-25			76.63	CD94; CHAIN: NULL;	NK CELL NK CELL, RECEPTOR, C-TYPE LECTIN, C-TYPE LECTIN-LIKE, NKD
950	1bj3	A	112	233	2.8e-36	0.21	0.69		COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A;	COLLAGEN BINDING PROTEIN IX-BP; IX-BP; COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									IX-BINDING PROTEIN B; CHAIN: B;	SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN
950	1bj3	A	112	234	2.8e-36			66.57	COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A; COAGULATION FACTOR IX-BINDING PROTEIN B; CHAIN: B;	COLLAGEN BINDING PROTEIN IX-BP, IX-BP, COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2 SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN
950	1c3a	B	110	236	1.4e-35			68.42	FLAVOCETIN-A: ALPHA SUBUNIT; CHAIN: A; FLAVOCETIN-A: BETA SUBUNIT; CHAIN: B	MEMBRANE PROTEIN C-TYPE LECTIN-LIKE DOMAINS
950	1c3a	B	112	236	1.4e-35	0.48	1.00		FLAVOCETIN-A: ALPHA SUBUNIT; CHAIN: A; FLAVOCETIN-A: BETA SUBUNIT; CHAIN: B	MEMBRANE PROTEIN C-TYPE LECTIN-LIKE DOMAINS
950	1c3a	B	1	117	1.3e-30			50.64	FLAVOCETIN-A: ALPHA SUBUNIT; CHAIN: A; FLAVOCETIN-A: BETA SUBUNIT; CHAIN: B	MEMBRANE PROTEIN C-TYPE LECTIN-LIKE DOMAINS
950	1dv8	A	113	233	1.4e-32	0.63	1.00		ASIALOGLYCOPROTEIN RECEPTOR 1; CHAIN: A;	SIGNALING PROTEIN HEPATIC LECTIN HI, C-TYPE LECTIN CRD
950	1dv8	A	113	241	1.4e-32			69.21	ASIALOGLYCOPROTEIN RECEPTOR 1; CHAIN: A;	SIGNALING PROTEIN HEPATIC LECTIN HI, C-TYPE LECTIN CRD
950	1e87	A	111	236	5.6e-27			81.55	EARLY ACTIVATION ANTIGEN CD69; CHAIN: A;	HEMATOPOIETIC CELL RECEPTOR ACTIVATION INDUCER MOLECULE (AIM), EA 1, HEMATOPOIETIC CELL RECEPTOR, LEUCOCYTE, C-TYPE LECTIN-LIKE, 2 NKD, KLR
950	1e87	A	1	117	1.1e-24			60.58	EARLY ACTIVATION ANTIGEN CD69; CHAIN:	HEMATOPOIETIC CELL RECEPTOR ACTIVATION

596

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A; INDUCER MOLECULE (AIM), EA 1, HEMATOPOIETIC CELL RECEPTOR, LEUCOCYTE, C-TYPE LECTIN-LIKE, 2 NKD, KIR	
950	legg	A	113	236	4.2e-28			53.59	MACROPHAGE MANNOSE RECEPTOR; CHAIN: A, B;	SUGAR BINDING PROTEIN C-TYPE LECTIN, MANNOSE RECEPTOR
950	legg	B	110	241	7e-29			60.69	MACROPHAGE MANNOSE RECEPTOR; CHAIN: A, B;	SUGAR BINDING PROTEIN C-TYPE LECTIN, MANNOSE RECEPTOR
950	lfvu	A	112	237	4.2e-30			52.01	BOTROCETIN ALPHA CHAIN: CHAIN: A, C; BOTROCETIN BETA CHAIN: CHAIN: B, D;	TOXIN PLATELET COAGGLUTININ ALPHA; PLATELET COAGGLUTININ BETA VON WILLEBRAND FACTOR MODULATOR, C-TYPE LECTIN, METAL-2 BINDING, LOOP EXCHANGED DIMER
950	lfvu	B	112	236	4.2e-34	0.58	0.99		BOTROCETIN ALPHA CHAIN: CHAIN: A, C; BOTROCETIN BETA CHAIN: CHAIN: B, D;	TOXIN PLATELET COAGGLUTININ ALPHA; PLATELET COAGGLUTININ BETA VON WILLEBRAND FACTOR MODULATOR, C-TYPE LECTIN, METAL-2 BINDING, LOOP EXCHANGED DIMER
950	lfvu	B	112	236	4.2e-34			69.59	BOTROCETIN ALPHA CHAIN: CHAIN: A, C; BOTROCETIN BETA CHAIN: CHAIN: B, D;	TOXIN PLATELET COAGGLUTININ ALPHA; PLATELET COAGGLUTININ BETA VON WILLEBRAND FACTOR MODULATOR, C-TYPE LECTIN, METAL-2 BINDING, LOOP EXCHANGED DIMER
950	lfvu	B	1	117	1.4e-28			50.65	BOTROCETIN ALPHA CHAIN: CHAIN: A, C; BOTROCETIN BETA CHAIN: CHAIN: B, D;	TOXIN PLATELET COAGGLUTININ ALPHA; PLATELET COAGGLUTININ BETA VON WILLEBRAND FACTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										MODULATOR, C-TYPE LECTIN, METAL-2 BINDING, LOOP EXCHANGED DIMER
950	1hq8	A	108	237	4.2e-30			93.99	NKG2-D; CHAIN: A;	APOPTOSIS HOMODIMER, CIS-PROLINE
950	1hq8	A	1	118	2.8e-28			66.38	NKG2-D; CHAIN: A;	APOPTOSIS HOMODIMER, CIS-PROLINE
950	1h9r	A	107	237	4.2e-28			92.55	NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-RELATED PROTEIN A; CHAIN: C;	IMMUNE SYSTEM NKG2D, MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-I, COMPLEX, IMMUNE SYSTEM
950	1h9r	A	1	118	5.6e-24			65.31	NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-RELATED PROTEIN A; CHAIN: C;	IMMUNE SYSTEM NKG2D, MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-I, COMPLEX, IMMUNE SYSTEM
950	1iod	A	112	233	1.3e-33	0.13	0.82		COAGULATION FACTOR X BINDING PROTEIN; CHAIN: A; COAGULATION FACTOR X BINDING PROTEIN; CHAIN: B; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G;	HYDROLASE/HYDROLASE INHIBITOR CALCIUM BRIDGING, DOMAIN SWAPPING
950	1iod	A	112	234	1.3e-33			59.17	COAGULATION FACTOR X BINDING PROTEIN; CHAIN: A; COAGULATION FACTOR X BINDING PROTEIN; CHAIN: B; COAGULATION FACTOR	HYDROLASE/HYDROLASE INHIBITOR CALCIUM BRIDGING, DOMAIN SWAPPING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
950	liod	B	112	236	9.8e-36	0.56	1.00		X GLA DOMAIN; CHAIN: G <sub>2</sub> COAGULATION FACTOR X BINDING PROTEIN; CHAIN: A; COAGULATION FACTOR X BINDING PROTEIN; CHAIN: B; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G <sub>2</sub>	HYDROLASE/HYDROLASE INHIBITOR CALCIUM BRIDGING, DOMAIN SWAPPING
950	liod	B	112	236	9.8e-36			60.25	COAGULATION FACTOR X BINDING PROTEIN; CHAIN: A; COAGULATION FACTOR X BINDING PROTEIN; CHAIN: B; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G <sub>2</sub>	HYDROLASE/HYDROLASE INHIBITOR CALCIUM BRIDGING, DOMAIN SWAPPING
950	lixx	A	112	233	7e-34	0.19	0.76		COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G <sub>2</sub>	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
950	lixx	A	112	234	7e-34			62.61	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G <sub>2</sub>	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
950	lixx	B	112	236	9.8e-36	0.48	1.00		COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G <sub>2</sub>	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER

599

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									C, D, E, F;	DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
950	1ixx	B	112	236	9.8e-36			65.36	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
950	1lit		113	235	4.2e-35	0.65	1.00		LITHOSTATHINE; CHAIN: NULL	PANCREATIC STONE INHIBITOR PANCREATIC STONE INHIBITOR, LECTIN
950	1lit		113	241	4.2e-35			59.42	LITHOSTATHINE; CHAIN: NULL	PANCREATIC STONE INHIBITOR PANCREATIC STONE INHIBITOR, LECTIN
950	1lit		1	122	1.3e-28			52.81	LITHOSTATHINE; CHAIN: NULL	PANCREATIC STONE INHIBITOR PANCREATIC STONE INHIBITOR, LECTIN
950	1qdd	A	101	235	9.8e-36	0.52	1.00		LITHOSTATHINE; CHAIN: A;	METAL BINDING PROTEIN PANCREATIC STONE PROTEIN, PSP; PANCREATIC STONE INHIBITOR, LITHOSTATHINE
950	1qdd	A	101	241	9.8e-36			64.36	LITHOSTATHINE; CHAIN: A;	METAL BINDING PROTEIN PANCREATIC STONE PROTEIN, PSP; PANCREATIC STONE INHIBITOR, LITHOSTATHINE
950	1qo3	C	105	236	6.8e-27			72.18	MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
950	1qo3	C	1	117	4.2e-25			52.76	LY49A; CHAIN: C, D; MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2- MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	LY49, LY-49 COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
950	1qo3	D	115	238	7e-26			65.45	MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2- MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
950	1qo3	D	1	119	4.2e-25			56.68	MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2- MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
950	1tm3		108	238	2.8e-28			50.25	TETRANECTIN; CHAIN: NULL;	LECTIN TETRANECTIN, PLASMINOGEN BINDING, KRINGLE 4, C-TYPE LECTIN, 2 CARBOHYDRATE RECOGNITION DOMAIN
950	2afp	A	107	241	2.8e-29			63.07	SEA RAVEN TYPE II ANTIFREEZE PROTEIN;	ANTIFREEZE PROTEIN RECOMBINANT SEA RAVEN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A;	PROTEIN, SOLUTION BACKBONE FOLD, C-2 TYPE LECTIN, ANTIFREEZE PROTEIN
951	1bj3	A	198	280	4.2e-26	0.10	0.09		COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A;	COLLAGEN BINDING PROTEIN IX-BP; IX-BP; COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2 SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN
									FLAVOCETIN-A: ALPHA SUBUNIT; CHAIN: A; FLAVOCETIN-A: BETA SUBUNIT; CHAIN: B	MEMBRANE PROTEIN C-TYPE LECTIN-LIKE DOMAINS
951	1c3a	B	197	281	4.2e-27	0.10	0.39		ASIALOGLYCOPROTEIN RECEPTOR 1; CHAIN: A; BOTROCECTIN ALPHA CHAIN; CHAIN: A, C; BOTROCECTIN BETA CHAIN; CHAIN: B, D;	SIGNALING PROTEIN HEPATIC LECTIN H; C-TYPE LECTIN CRD TOXIN PLATELET COAGELUTININ ALPHA; PLATELET COAGELUTININ BETA VON WILBRAND FACTOR MODULATOR, C-TYPE LECTIN, METAL-2 BINDING, LOOP EXCHANGED DIMER
951	1dv8	A	198	276	1.4e-24	0.22	0.68		NKG2-D; CHAIN: A;	APOPTOSIS HOMODIMER, CIS-PROLINE
951	1fva	B	198	280	1.4e-26	0.07	0.47		NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-RELATED PROTEIN A; CHAIN: C;	IMMUNE SYSTEM NKG2D, MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-1, COMPLEX, IMMUNE SYSTEM
									COAGULATION FACTOR X BINDING PROTEIN; CHAIN: A;	HYDROLASE/HYDROLASE INHIBITOR CALCIUM BRIDGING, DOMAIN SWAPPING
951	1iod	A	198	280	7e-25	0.02	0.19			

602

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
951	1iod	B	198	280	2.8e-25	-0.02	0.01		COAGULATION FACTOR X BINDING PROTEIN; CHAIN: B; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G;	HYDROLASE/HYDROLASE INHIBITOR CALCIUM BRIDGING, DOMAIN SWAPPING
951	1ixx	B	198	281	4.2e-26	-0.12	0.09		COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F; CHAIN: B; COAGULATION FACTOR X GLA DOMAIN; CHAIN: G;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
951	2afp	A	193	281	1.4e-24	0.38	0.46		SEA RAVEN TYPE II ANTIFREEZE PROTEIN; CHAIN: A;	ANTIFREEZE PROTEIN RECOMBINANT SEA RAVEN PROTEIN, SOLUTION BACKBONE FOLD, C-2 TYPE LECTIN, ANTIFREEZE PROTEIN
952	1g9r	A	1242	1507	1e-44	0.19	0.93		GLYCOSYL TRANSFERASE; CHAIN: A;	TRANSFERASE ALPHA-BETA STRUCTURE
953	1al7		68	229	2.8e-16	0.33	1.00		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE

603

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
953	1elr	A	152	227	0.00014	0.14	0.60		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
953	1elr	A	65	218	1e-09	0.22	0.99		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
953	1elw	A	65	217	1.7e-13	0.47	0.81		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
953	1elw	A	68	215	1.3e-08	0.44	0.95		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
953	1fch	A	71	226	3.4e-09	0.20	0.62		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISOMAL RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
953	1hh8	A	150	198	2.8e-05	0.06	0.43		NEUTROPHIL CYTOSOL FACTOR 2; CHAIN: A;	PHAGOCYTE OXIDASE FACTOR P67PHOX, NCF-2; PHAGOCYTE OXIDASE FACTOR, SH3 DOMAIN, REPEAT, TPR REPEAT
953	1ihg	A	68	226	8.4e-37	0.39	1.00		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PIASE IMMUNOPHILIN TETRATRICOPEPTIDE
957	1fnf		1	203	5.6e-32	-0.35	0.18		FIBRONECTIN; IFNF 6 CHAIN: NULL; IFNF 7	CELL ADHESION PROTEIN RGD, EXTRACELLULAR MATRIX IFNF 18
957	1fnh	A	1	192	1.3e-27	-0.20	0.11		FIBRONECTIN; CHAIN:	HEPARIN AND INTEGRIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A; FIBRONECTIN; CHAIN: A; B	BINDING HEPARIN AND INTEGRIN BINDING
957	1fth	A	31	201	2.8e-22	-0.22	0.27			HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
957	1ilr	A	29	202	1.3e-12	0.14	-0.19		INTERLEUKIN-6 RECEPTOR BETA CHAIN; CHAIN: A; VIRAL IL-6; CHAIN: B;	CYTOKINE GP130; FUNCTIONAL INTERLEUKIN-6 HOMOLOG; CYTOKINE/RECEPTOR COMPLEX, GP130, VIRAL IL-6, CRYSTAL 2 STRUCTURE
957	1mfh		29	204	4.2e-29			66.80	FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN
957	1mfh		2	112	1.4e-23	0.12	-0.08		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN
957	1mfh		31	203	4.2e-29	0.11	0.41		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN
957	1qg3	A	29	194	4.2e-24			57.81	INTEGRIN BETA-4 SUBUNIT; CHAIN: A, B;	STRUCTURAL PROTEIN INTEGRIN, HEMIDESMOSOME, FIBRONECTIN, CARCINOMA, STRUCTURAL 2 PROTEIN
957	1qt4	A	31	204	2.8e-21			71.24	TENASCIN; CHAIN: A, B;	STRUCTURAL PROTEIN TENASCIN, FIBRONECTIN TYPE-III, HEPARIN, EXTRACELLULAR 2 MATRIX, ADHESION, FUSION PROTEIN, STRUCTURAL PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
957	1qr4	A	33	193	2.8e-21	0.01	0.27		TENASCIN; CHAIN: A, B;	STRUCTURAL PROTEIN TENASCIN, FIBRONECTIN TYPE-III, HEPARIN, EXTRACELLULAR 2 MATRIX, ADHESION, FUSION PROTEIN, STRUCTURAL PROTEIN
957	1qr4	A	3	117	7e-20	-0.25	0.84		TENASCIN; CHAIN: A, B;	STRUCTURAL PROTEIN TENASCIN, FIBRONECTIN TYPE-III, HEPARIN, EXTRACELLULAR 2 MATRIX, ADHESION, FUSION PROTEIN, STRUCTURAL PROTEIN
957	1ten		11	98	8.4e-16			53.41	CELL ADHESION PROTEIN TENASCIN (THIRD FIBRONECTIN TYPE III REPEAT) 1TEN 3	
957	1ten		30	117	1.7e-18			53.70	CELL ADHESION PROTEIN TENASCIN (THIRD FIBRONECTIN TYPE III REPEAT) 1TEN 3	
957	1tff		11	99	4.2e-21			51.26	GLYCOPROTEIN FIBRONECTIN (TENTH TYPE III MODULE) (NMR, 36 STRUCTURES) 1TTF 3	
957	1tff		30	121	8.4e-18			51.78	GLYCOPROTEIN FIBRONECTIN (TENTH TYPE III MODULE) (NMR, 36 STRUCTURES) 1TTF 3	
958	1oge		123	282	5.6e-71	1.17	1.00		HYDROLASE (METALLOPROTEASE) COLLAGENASE (E.C.3.4.24.7) (CATALYTIC DOMAIN) CRYSTAL FORM I 10GB 3	

606

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
958	1cge		123	283	5.1e-74			280.54	HYDROLASE (METALLOPROTEASE) COLLAGENASE (E.C.3.4.24.7) (CATALYTIC DOMAIN) CRYSTAL FORM I 1CGE 3	
958	1cge		123	284	5.1e-74	1.12	1.00		HYDROLASE (METALLOPROTEASE) COLLAGENASE (E.C.3.4.24.7) (CATALYTIC DOMAIN) CRYSTAL FORM I 1CGE 3	
958	1cge		76	188	4.2e-55			176.37	HYDROLASE (METALLOPROTEASE) COLLAGENASE (E.C.3.4.24.7) (CATALYTIC DOMAIN) CRYSTAL FORM I 1CGE 3	
958	1ck7	A	27	246	4.2e-73	0.27	0.99		GELATINASE A; CHAIN: A; 3	HYDROLASE MMP-2, 72KD TYPE IV COLLAGENASE; HYDROLASE (METALLOPROTEASE), FULL- LENGTH, METALLOPROTEINASE, 2 GELATINASE A
958	1fb1		121	487	0	1.03	1.00		FIBROBLAST (INTERSTITIAL) COLLAGENASE (MMP- 1); 1FBL 4 CHAIN: NULL; 1FBL 5	METALLOPROTEASE
958	1fb1		121	487	0			592.38	FIBROBLAST (INTERSTITIAL) COLLAGENASE (MMP- 1); 1FBL 4 CHAIN: NULL;	METALLOPROTEASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
958	1fbl		74	188	7e-58			57.48	1FBL 5 FIBROBLAST (INTERSTITIAL) COLLAGENASE (MMP-1); 1FBL 4 CHAIN: NULL; 1FBL 5	METALLOPROTEASE
958	1hfc		128	282	1.4e-70	1.08	1.00		METALLOPROTEASE FIBROBLAST COLLAGENASE (E.C.3.4.24.7) 1HFC 3	
958	1hfc		128	284	1.4e-70			274.25	METALLOPROTEASE FIBROBLAST COLLAGENASE (E.C.3.4.24.7) 1HFC 3	
958	1hfc		81	188	5.6e-55			168.69	METALLOPROTEASE FIBROBLAST COLLAGENASE (E.C.3.4.24.7) 1HFC 3	
958	1lbu		50	127	0.0014	0.42	0.24		MURAMOYL- PENTAPEPTIDE CARBOXYPEPTIDASE; CHAIN: NULL;	HYDROLASE HYDROLASE, NUCLEAR RECEPTOR, CARBOXYPEPTIDASE
958	1sat		47	490	1.4e-34			78.13	SERRATIA PROTEASE ISAT 3	HYDROLASE (SERINE PROTEASE) SERRALYSIN, MAJOR METALLO PROTEASE FROM ISAT 4 PARALLEL BETA HELIX, PARALLEL BETA ROLL ISAT 7
958	1sat		83	406	1.4e-34	0.21	0.39		SERRATIA PROTEASE ISAT 3	HYDROLASE (SERINE PROTEASE) SERRALYSIN, MAJOR METALLO PROTEASE FROM ISAT 4 PARALLEL BETA HELIX, PARALLEL BETA ROLL ISAT 7
958	1slm		54	285	1.1e-96	0.76	1.00		STROMELYSIN-1; CHAIN: NULL;	HYDROLASE MATRIX METALLOPROTEINASE-3;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										PROTEOGLYCANASE; HYDROLASE, METALLOPROTEASE, FIBROBLAST, COLLAGEN DEGRADATION
958	1slm		55	285	1.1e-96			271.16	STROMELYSIN-1; CHAIN: NULL;	HYDROLASE MATRIX METALLOPROTEINASE-3, PROTEOGLYCANASE; HYDROLASE, METALLOPROTEASE, FIBROBLAST, COLLAGEN DEGRADATION
958	1slm		8	187	2.8e-83			190.20	STROMELYSIN-1; CHAIN: NULL;	HYDROLASE MATRIX METALLOPROTEINASE-3, PROTEOGLYCANASE; HYDROLASE, METALLOPROTEASE, FIBROBLAST, COLLAGEN DEGRADATION
962	1ap0		35	106	3e-14			50.89	MODIFIER PROTEIN 1; CHAIN: NULL;	CHROMATIN-BINDING MOMOD1, HETEROCHROMATIN PROTEIN 1; CHROMATIN-BINDING, PROTEIN INTERACTION MOTIF, ALPHA+BETA
964	1a0q	L	20	242	8.4e-66			110.96	29G11 FAB; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, ESTERASE
964	1a2y	A	1	101	9.8e-39			51.49	MONOCLONAL ANTIBODY D1.3; CHAIN: A, B; LYSOZYME, CHAIN: C;	COMPLEX (IMMUNOGLOBULIN/HYDROLAS E) COMPLEX (IMMUNOGLOBULIN/HYDROLAS E), IMMUNOGLOBULIN V 2 REGION, SIGNAL, HYDROLASE, GLYCOSIDASE, BACTERIOLYTIC 3 ENZYME, EGG WHITE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
964	1a7q	L	1	101	4.2e-37			53.02	MONOCLONAL ANTIBODY D1.3; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, VARIANT
964	1adq	L	21	246	8.4e-72			111.15	IGG4 REA; CHAIN: A; RF-ANTIGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
964	1ac6	L	19	243	2.8e-65			107.14	ANTIBODY CTM01; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB FRAGMENT, HUMANISATION
964	1afv	H	24	239	5.6e-85	0.22	-1202.08		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN) HIV-1 CA, HIV CA, HIV P24, P24; FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN), HIV, CAPSID PROTEIN, 2 P24
964	1ahw	A	19	243	2.8e-71			109.26	IMMUNOGLOBULIN FAB 5G9; CHAIN: A, B, D, E; TISSUE FACTOR; CHAIN: C, F;	COMPLEX (IMMUNOGLOBULIN/TISSUE FACTOR) FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN; TF, THROMBOPLASTIN, COAGULATION FACTOR III; BLOOD COAGULATION, TISSUE FACTOR, FAB, COMPLEX, ANTIBODY, 2 COMPLEX (IMMUNOGLOBULIN/TISSUE FACTOR)
964	1aif	A	19	245	5.6e-69			107.73	ANTI-IDIOTYPIC FAB 409.5.3 (IGG2A) FAB; CHAIN: A, B, L, H	IMMUNOGLOBULIN IMMUNOGLOBULIN, C REGION, V REGION
964	1aqk	L	21	246	2.8e-68			110.74	FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
964	1ay1	L	19	238	7e-72			111.56	TP7 FAB; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, ANTIBODY, FAB, ENZYME INHIBITOR, PCR, 2 HOT START
964	1b0w	A	1	101	4.2e-45			50.86	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C;	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
964	1ba1	L	19	243	1.4e-71			107.96	IMMUNOGLOBULIN FAB FRAGMENT OF MURINE MONOCLONAL ANTIBODY AN02 COMPLEX IBAF 3 WITH ITS HAPTEN (2,2,6,6-TETRAMETHYL-1-PIPERIDINYLOXY - IBAF 4 DINITROPHENYL) IBAF 5	
964	1bj1	J	19	242	2.8e-74			107.19	FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
964	1bjm	A	20	246	1.4e-66			115.09	LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER; IBJM 6 CHAIN: A, B; IBJM 7	IMMUNOGLOBULIN BENCE-JONES PROTEIN; IBJM 8 BENCE JONES, ANTIBODY, MULTIPLE QUATERNARY STRUCTURES IBJM 13
964	1bvk	A	1	101	1.1e-44			52.15	HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
964	1bw w	A	1	101	4.2e-47			50.82	IG KAPPA CHAIN V-L REGION RE; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENGE-JONES 2 PROTEIN, IMMUNE SYSTEM
964	1c1g	A	1	234	8.4e-43			63.76	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
964	1d5i	L	19	238	9.8e-69			108.44	CHIMERIC GERMLINE PRECURSOR OF OXY-COPE CHAIN: L; CHIMERIC GERMLINE PRECURSOR OF OXY-COPE CHAIN: H;	IMMUNE SYSTEM IMMUNE SYSTEM
964	1e4x	H	24	238	1.3e-81	0.14	-1202.08		TAB2; CHAIN: L, M; TAB2; CHAIN: H, I; CYCLIC PEPTIDE; CHAIN: P, Q	COMPLEX (ANTIBODY/ANTIGEN) CROSS-REACTIVITY, PROTEIN-PEPTIDE RECOGNITION
964	1e6o	H	24	239	5.6e-82	0.22	-1202.08		IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L; IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H;	IMMUNOGLOBULIN FAB, ANTIBODY, ANTIGEN, HIV-1, P24, CA
964	1f8t	L	19	243	2.8e-65			107.82	ANTIBODY FAB FRAGMENT (LIGHT CHAIN); CHAIN: L; ANTIBODY FAB FRAGMENT (HEAVY CHAIN); CHAIN: H	IMMUNE SYSTEM MONOCLONAL ANTIBODY, ANTIGEN-BINDING FRAGMENT, INTERLEUKIN-22, X-RAY ANALYSIS, CRYSTAL
964	1fbi	H	24	239	2.8e-81	0.07	-1202.08		COMPLEX	

612

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(ANTIBODY/ANTIGEN) FAB FRAGMENT OF THE MONOCLONAL ANTIBODY F9.13.7 (IGG1) IFBI 3 COMPLEXED WITH LYSOZYME (E.C.3.2.1.17) IFBI 4	
964	1f9n	H	24	238	2.8e-83	0.25	-1202.08		IMMUNOGLOBULIN FAB 5G9; CHAIN: L, H;	IMMUNOGLOBULIN FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN; ANTIBODY, FAB, ANTI-TF, MONOCLONAL, MURINE, IMMUNOGLOBULIN
964	1fj1	A	19	242	5.6e-68			113.25	HYBRIDOMA ANTIBODY LA2 (LIGHT CHAIN); CHAIN: A, C; HYBRIDOMA ANTIBODY LA2 (HEAVY CHAIN); CHAIN: B, D; OUTER SURFACE PROTEIN A; CHAIN: E, F;	IMMUNE SYSTEM LA2 FAB; LA2 FAB; OSP A; OSP A, LYME DISEASE, ANTIBODY FAB FRAGMENT, NEUTRALIZING 2 EPTOPE
964	1fhs	L	19	243	9.8e-71			107.73	IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
964	1hi6	A	19	243	1.3e-66			109.95	IGG2A KAPPA ANTIBODY CB41 (LIGHT CHAIN); CHAIN: A; IGG2A KAPPA ANTIBODY CB41 (HEAVY CHAIN); CHAIN: B; PEPTIDE 5; CHAIN: C;	COMPLEX (ANTIBODY/PEPTIDE) POLYSPECIFICITY, CROSSREACTIVITY, FAB-FRAGMENT, PEPTIDE, 2 HIV-1

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
964	1hxm	B	19	243	9e-57			242.19	GAMMA-DELTA T-CELL RECEPTOR; CHAIN: A, C, E, G; GAMMA-DELTA T-CELL RECEPTOR; CHAIN: B, D, F, H;	IMMUNE SYSTEM T-CELL RECEPTOR DELTA CHAIN; T-CELL RECEPTOR GAMMA CHAIN; IG DOMAIN, T CELL RECEPTOR, TCR, GDTOR
964	1hxm	B	1	84	8.4e-34			50.67	GAMMA-DELTA T-CELL RECEPTOR; CHAIN: A, C, E, G; GAMMA-DELTA T-CELL RECEPTOR; CHAIN: B, D, F, H;	IMMUNE SYSTEM T-CELL RECEPTOR DELTA CHAIN; T-CELL RECEPTOR GAMMA CHAIN; IG DOMAIN, T CELL RECEPTOR, TCR, GDTOR
964	1hzh	H	25	296	2.8e-83	0.11	-1202.08		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
964	1igt	B	145	296	1.4e-12	0.14	-1202.08		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
964	1igt	B	24	283	8.4e-90	0.06	-1202.08		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
964	1kb5	L	19	243	2.8e-70			109.47	KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2
964	1hl	A	22	246	4.2e-69			114.77	LAMBDA III BENGE JONES PROTEIN CLE; CHAIN: A, B	(IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN IMMUNOGLOBULIN, BENGE JONES PROTEIN
964	1ngp	H	24	239	2.8e-82	0.09	-1202.08		NI9 (GGI=LAMBDA=); CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN,

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
964	1wtl	A	1	101	4.2e-45			50.35	IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL 3 (BENCE-JONES PROTEIN) 1WTL 4	
964	25c8	H	24	239	4.2e-83	0.12	-1202.08		IGG 5C8; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
964	2fb4	L	20	246	4.2e-67			109.74	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB 2FB4 4	
964	2mcg	1	20	246	9.8e-69			108.78	IMMUNOGLOBULIN IMMUNOGLOBULIN LAMBDA LIGHT CHAIN DIMER ((MCG\$) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
964	8fab	A	23	241	1.4e-69			118.67	IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
968	1cdo	A	157	495	1.4e-78			91.38	ALCOHOL DEHYDROGENASE; ICDO 6 CHAIN: A, B; ICDO 7	OXIDOREDUCTASE (CH-OH(D)-NAD(A)) OXIDOREDUCTASE ICDO 15
968	1cdo	A	164	492	1.4e-78	0.33	-1202.08		ALCOHOL DEHYDROGENASE; ICDO 6 CHAIN: A, B; ICDO 7	OXIDOREDUCTASE (CH-OH(D)-NAD(A)) OXIDOREDUCTASE ICDO 15
968	1cdo	A	1	327	2.8e-69			82.11	ALCOHOL	OXIDOREDUCTASE (CH-OH(D)-

Table 5

SEQ ID No.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DEHYDROGENASE; ICDO 6 CHAIN: A, B; ICDO 7	NAD(A)) OXIDOREDUCTASE ICDO 15
968	Idlt	A	157	495	1.1e-80			92.65	ALCOHOL DEHYDROGENASE CLASS IV SIGMA CHAIN; CHAIN: A, B, C, D;	OXIDOREDUCTASE RETINOL DEHYDROGENASE, GASTRIC ALCOHOL, ROSSMANN OR DINUCL EOTIDE FOLD
968	Idlt	A	164	492	1.1e-80	0.32	-1202.08		ALCOHOL DEHYDROGENASE CLASS IV SIGMA CHAIN; CHAIN: A, B, C, D;	OXIDOREDUCTASE RETINOL DEHYDROGENASE, GASTRIC ALCOHOL, ROSSMANN OR DINUCL EOTIDE FOLD
968	Idlt	A	3	327	5.6e-72			81.94	ALCOHOL DEHYDROGENASE CLASS IV SIGMA CHAIN; CHAIN: A, B, C, D;	OXIDOREDUCTASE RETINOL DEHYDROGENASE, GASTRIC ALCOHOL, ROSSMANN OR DINUCL EOTIDE FOLD
968	Idet	A	110	495	8.4e-87			100.84	HUMAN BETAI ALCOHOL DEHYDROGENASE; IDEH 7 CHAIN: A, B; IDEH 8	OXIDOREDUCTASE BETAI ADH; IDEH 9 NAD+ DEPENDENT ALCOHOL, DEHYDROGENASE IDEH 26
968	Idet	A	164	492	8.4e-87	0.41	-1202.08		HUMAN BETAI ALCOHOL DEHYDROGENASE; IDEH 7 CHAIN: A, B; IDEH 8	OXIDOREDUCTASE BETAI ADH; IDEH 9 NAD+ DEPENDENT ALCOHOL, DEHYDROGENASE IDEH 26
968	Idet	A	1	327	1.4e-78			87.63	HUMAN BETAI ALCOHOL DEHYDROGENASE; IDEH 7 CHAIN: A, B; IDEH 8	OXIDOREDUCTASE BETAI ADH; IDEH 9 NAD+ DEPENDENT ALCOHOL, DEHYDROGENASE IDEH 26
968	Idet	A	106	495	7e-74			100.96	ALCOHOL DEHYDROGENASE, CLASS II; CHAIN: A, B;	ALCOHOL, DEHYDROGENASE ALCOHOL, DEHYDROGENASE
968	Idet	A	164	492	7e-74	0.38	-1202.08		ALCOHOL	ALCOHOL, DEHYDROGENASE

Table 5

SEQ ID No:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DEHYDROGENASE, CLASS II; CHAIN: A, B;	ALCOHOL DEHYDROGENASE
968	1e3i	A	3	327	5.6e-67			82.29	ALCOHOL DEHYDROGENASE, CLASS II; CHAIN: A, B;	ALCOHOL DEHYDROGENASE
968	1e3j	A	163	491	1.4e-60			82.26	NADP(H)-DEPENDENT KETOSE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE
968	1e3j	A	1	323	2.8e-58			78.55	NADP(H)-DEPENDENT KETOSE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE
968	1het	A	110	495	4.2e-83			94.36	ALCOHOL DEHYDROGENASE E CHAIN; CHAIN: A, B;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE I; OXIDOREDUCTASE(NAD(A)-CHOH(D))
968	1het	A	164	492	4.2e-83	0.46	-1202.08		ALCOHOL DEHYDROGENASE E CHAIN; CHAIN: A, B;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE I; OXIDOREDUCTASE(NAD(A)-CHOH(D))
968	1het	A	3	327	2.8e-75			79.40	ALCOHOL DEHYDROGENASE E CHAIN; CHAIN: A, B;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE I; OXIDOREDUCTASE(NAD(A)-CHOH(D))
968	1hso	A	110	495	7e-87			101.16	CLASS I ALCOHOL DEHYDROGENASE I, ALPHA SUBUNIT; CHAIN: A, B;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE (CLASS I), ALPHA ROSSMANN FOLD, ALCOHOL DEHYDROGENASE
968	1hso	A	164	492	7e-87	0.46	-1202.08		CLASS I ALCOHOL DEHYDROGENASE I, ALPHA SUBUNIT; CHAIN: A, B;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE (CLASS I), ALPHA ROSSMANN FOLD, ALCOHOL DEHYDROGENASE
968	1hso	A	1	327	1.4e-78			89.46	CLASS I ALCOHOL DEHYDROGENASE I, ALPHA SUBUNIT;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE (CLASS I), ALPHA ROSSMANN FOLD,



617

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
968	1ht0	A	110	495	4.2e-86			105.32	CHAIN: A, B; CLASS I ALCOHOL DEHYDROGENASE 3, GAMMA SUBUNIT; CHAIN: A, B;	ALCOHOL DEHYDROGENASE OXIDOREDUCTASE ALCOHOL DEHYDROGENASE (CLASS I), GAMMA ROSSMANN FOLD, ALCOHOL DEHYDROGENASE, ZINC
968	1ht0	A	164	492	4.2e-86	0.33	-1202.08		CLASS I ALCOHOL DEHYDROGENASE 3, GAMMA SUBUNIT; CHAIN: A, B;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE (CLASS I), GAMMA ROSSMANN FOLD, ALCOHOL DEHYDROGENASE, ZINC
968	1ht0	A	3	327	4.2e-78			90.30	CLASS I ALCOHOL DEHYDROGENASE 3, GAMMA SUBUNIT; CHAIN: A, B;	OXIDOREDUCTASE ALCOHOL DEHYDROGENASE (CLASS I), GAMMA ROSSMANN FOLD, ALCOHOL DEHYDROGENASE, ZINC
968	1kev	A	1	325	9.8e-48			66.86	NADP-DEPENDENT ALCOHOL DEHYDROGENASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE OXIDOREDUCTASE, ZINC, NADP
968	1qor	A	165	495	2.8e-77			132.05	OXIDOREDUCTASE QUINONE OXIDOREDUCTASE COMPLEXED WITH NADPH 1QOR 3	
968	1qor	A	168	493	2.8e-77	0.62	-1202.08		OXIDOREDUCTASE QUINONE OXIDOREDUCTASE COMPLEXED WITH NADPH 1QOR 3	
968	1qor	A	2	327	5.6e-78			128.24	OXIDOREDUCTASE QUINONE OXIDOREDUCTASE COMPLEXED WITH	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
968	1tcf		11	162	0.008			52.38	NADPH 100R 3 TROPONIN C; CHAIN: NUL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
968	1teh	A	112	495	1.4e-90			100.75	HUMAN CHICH1 ALCOHOL DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE GLUTATHIONE-DEPENDENT FORMALDEHYDE DEHYDROGENASE, NAD+ DEPENDENT ALCOHOL DEHYDROGENASE 2 GLUTATHIONE DEPENDENT FORMALDEHYDE DEHYDROGENASE
968	1teh	A	164	492	1.4e-90	0.37	-1202.08		HUMAN CHICH1 ALCOHOL DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE GLUTATHIONE-DEPENDENT FORMALDEHYDE DEHYDROGENASE, NAD+ DEPENDENT ALCOHOL DEHYDROGENASE 2 GLUTATHIONE DEPENDENT FORMALDEHYDE DEHYDROGENASE
968	1teh	A	3	327	4.2e-78			85.11	HUMAN CHICH1 ALCOHOL DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE GLUTATHIONE-DEPENDENT FORMALDEHYDE DEHYDROGENASE, NAD+ DEPENDENT ALCOHOL DEHYDROGENASE 2 GLUTATHIONE DEPENDENT FORMALDEHYDE

619

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
968	1ykf	A	165	494	2.8e-58			68.99	NADP-DEPENDENT ALCOHOL DEHYDROGENASE; CHAIN: A, B, C, D;	DEHYDROGENASE
968	1ykf	A	1	326	1.4e-49			67.17	NADP-DEPENDENT ALCOHOL DEHYDROGENASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE
968	7adh		1	327	5.6e-60			62.04	OXIDOREDUCTASE (NAD(A)-CHOH(D)) ISONICOTINIMIDYLATE D LIVER ALCOHOL DEHYDROGENASE 7ADH 4 (E.C.1.1.1.1) 7ADHD 1 7ADHD 2	
970	1bfd		14	349	4.2e-39	0.14	-1202.08		BENZOYLFORMATE DECARBOXYLASE; CHAIN: NULL;	LYASE LYASE, CARBON-CARBON, DECARBOXYLASE, MANDELATE CATABOLISM, 2 THIAMIN DIPHOSPHATE
970	1dtw	A	1	316	8.4e-70			118.95	BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE CHAIN: A; BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE CHAIN: B;	OXIDOREDUCTASE THDP-BINDING FOLD, BRANCHED-CHAIN ALPHA-KETO ACID 2 DEHYDROGENASE
970	1dtw	A	73	453	5.6e-70			161.96	BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE CHAIN: A; BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE CHAIN: B;	OXIDOREDUCTASE THDP-BINDING FOLD, BRANCHED-CHAIN ALPHA-KETO ACID 2 DEHYDROGENASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
970	1dtw	A	77	449	5.6e-70	0.44	-1202.08		BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE CHAIN: A; BRANCHED-CHAIN ALPHA-KETO ACID DEHYDROGENASE CHAIN: B;	OXIDOREDUCTASE THDP-BINDING FOLD, BRANCHED-CHAIN ALPHA-KETO ACID 2 DEHYDROGENASE
970	1pox	A	15	392	7e-40	0.22	-1202.08		OXIDOREDUCTASE(OXYGEN AS ACCEPTOR) PYRUVATE OXIDASE (E.C.1.2.3.3) MUTANT WITH PRO 178 IPOX 3 REPLACED BY SER, SER 188 REPLACED BY ASN, AND ALA 458 IPOX 4 REPLACED BY VAL (P178S,S188N,A458V) IPOX 5	
970	1pyd	A	15	365	1.3e-46	0.06	-1202.08		LYASE (CARBON-CARBON) PYRUVATE DECARBOXYLASE (PDC) (E.C.4.1.1.1) IPVD 3	
970	1qso	A	121	446	2.8e-59	0.61	-1202.08		2-OXOISOVALERATE DEHYDROGENASE ALPHA-SUBUNIT; CHAIN: A; 2-OXOISOVALERATE DEHYDROGENASE BETA-SUBUNIT; CHAIN: B;	OXIDOREDUCTASE HETEROTETRAMER, THDP COFACTOR, OXIDOREDUCTASE
970	1qso	A	1	317	2.8e-59			91.32	2-OXOISOVALERATE DEHYDROGENASE ALPHA-SUBUNIT; CHAIN: A; 2-OXOISOVALERATE	OXIDOREDUCTASE HETEROTETRAMER, THDP COFACTOR, OXIDOREDUCTASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
970	1qs0	A	40	446	9e-76			140.12	DEHYDROGENASE BETA-SUBUNIT; CHAIN: B;	OXIDOREDUCTASE HETEROTETRAMER, THDP COFACTOR, OXIDOREDUCTASE
970	1qs0	A	68	445	9e-76	0.55	-1202.08		2-OXOISOVALERATE DEHYDROGENASE ALPHA-SUBUNIT; CHAIN: A; 2- OXOISOVALERATE DEHYDROGENASE BETA-SUBUNIT; CHAIN: B;	OXIDOREDUCTASE HETEROTETRAMER, THDP COFACTOR, OXIDOREDUCTASE
970	1tk	A	127	446	3e-57	0.16	-1202.08		TRANSFERRASE(KETONE RESIDUES) TRANSKETOLASE (E.C.2.2.1.1) 1TRK 3	
971	1a4j	A	24	139	1e-08	0.14	-1202.08		IMMUNOGLOBULIN, DIELS ALDER CATALYTIC ANTIBODY; CHAIN: L, H, A, B;	IMMUNOGLOBULIN ANTIBODY, CATALYTIC ANTIBODY, DIELS ALDER, 2 GERMLINE
971	1bd2	E	32	180	1.4e-67	0.07	-1202.08		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
971	1bd2	E	32	180	1.4e-67			86.31	CHAIN: E; HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)
971	1bec		32	180	5.6e-66			81.96	14.3.D T CELL ANTIGEN RECEPTOR; IBEC.5 CHAIN: NULL; IBEC 6	RECEPTOR T CELL RECEPTOR IBEC 14
971	1d9k	B	1	96	8.4e-38			57.60	T-CELL RECEPTOR D10 (ALPHA CHAIN); CHAIN: A, E; T-CELL RECEPTOR D10 (BETA CHAIN); CHAIN: B, F; MHC I-AK A CHAIN (ALPHA CHAIN); CHAIN: C, G; MHC I-AK B CHAIN (BETA CHAIN); CHAIN: D, H; CONALBUMIN PEPTIDE; CHAIN: P, Q;	IMMUNE SYSTEM MHC I-AK; MHC I-AK; T-CELL RECEPTOR, MHC CLASS II, D10, I-AK
971	1fyt	E	32	180	5.6e-61			81.57	HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR CHAIN: A; HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR-1 CHAIN: B; HEMAGGLUTININ HA1 PEPTIDE CHAIN; CHAIN: C; T-CELL RECEPTOR ALPHA CHAIN; CHAIN: D; T-	IMMUNE SYSTEM HLA-DR1, DR4; HLA-DR1, DRB1 0101; TCR HA1.7 ALPHA CHAIN; TCR HA1.7 BETA CHAIN; PROTEIN-PROTEIN COMPLEX, IMMUNOGLOBULIN FOLD

623

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
971	Infj	B	30	180	7e-64	0.19	-1202.08		CELL RECEPTOR BETA CHAIN; CHAIN: E;	
										COMPLEX (IMMUNORECEPTOR/IMMUNOGL OBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGL OBULIN)
971	Infj	B	30	180	7e-64			124.34	N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGL OBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGL OBULIN)
971	Iter	B	30	180	1.3e-64	0.11	-1202.08		ALPHA, BETA T-CELL RECEPTOR CHAIN: A, B;	RECEPTOR TCR; T-CELL, RECEPTOR, TRANSMEMBRANE, GLYCOPROTEIN, SIGNAL
971	Iter	B	30	180	1.3e-64			73.51	ALPHA, BETA T-CELL RECEPTOR CHAIN: A, B;	RECEPTOR TCR; T-CELL, RECEPTOR, TRANSMEMBRANE, GLYCOPROTEIN, SIGNAL
972	Icun	A	21	239	0.003			62.32	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2.2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
975	Iej	A	163	485	1.4e-18	0.07	-1202.08		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
975	Iej	A	182	532	9.8e-23			85.09	TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
975	Iej	A	212	527	9.8e-23	0.18	-1202.08		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
975	Iej	A	410	580	2.8e-19	0.02	-1202.08		TRANSCRIPTIONAL REPRESSOR TUP1;	TRANSCRIPTION INHIBITOR BETA-PROPELLER

624

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
975	1got	B	227	567	1.4e-20			85.50	CHAIN: A, B, C; GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT- GAMMA; CHAIN: G;	COMPLEX (GTP- BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP- BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
975	1got	B	305	579	1.4e-20	0.06	-1202.08		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT- GAMMA; CHAIN: G;	COMPLEX (GTP- BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP- BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
977	1evu	A	108	710	9.8e-91			142.67	PROTAGLANDIN H2 SYNTHASE-2; CHAIN: A, B; 9-MER; CHAIN: F;	OXIDOREDUCTASE COX-2, CYCLOOXYGENASE, PROTAGLANDIN, ARACHIDONATE, 2 ENDOPEROXIDE
977	1evu	A	155	631	9.8e-91	0.02	-1202.08		PROTAGLANDIN H2 SYNTHASE-2; CHAIN: A, B; 9-MER; CHAIN: F;	OXIDOREDUCTASE COX-2, CYCLOOXYGENASE, PROTAGLANDIN, ARACHIDONATE, 2 ENDOPEROXIDE
977	1evu	A	6	584	0			142.21	PROTAGLANDIN H2 SYNTHASE-2; CHAIN: A, B; 9-MER; CHAIN: F;	OXIDOREDUCTASE COX-2, CYCLOOXYGENASE, PROTAGLANDIN, ARACHIDONATE, 2 ENDOPEROXIDE
977	1d2v	A	132	235	1.5e-42			141.84	MYELOPEROXIDASE; CHAIN: A, B; MYELOPEROXIDASE;	OXIDOREDUCTASE HEME- PROTEIN, PEROXIDASE, OXIDOREDUCTASE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: C, D;	PEROXIDASE-2 BROMIDE COMPLEX
977	1d2v	A	6	109	2.8e-42			141.89	MYELOPEROXIDASE; CHAIN: A, B; MYELOPEROXIDASE; CHAIN: C, D;	OXIDOREDUCTASE HEME-PROTEIN, PEROXIDASE, OXIDOREDUCTASE, PEROXIDASE-2 BROMIDE COMPLEX
977	1d2v	C	244	733	0			561.59	MYELOPEROXIDASE; CHAIN: A, B; MYELOPEROXIDASE; CHAIN: C, D;	OXIDOREDUCTASE HEME-PROTEIN, PEROXIDASE, OXIDOREDUCTASE, PEROXIDASE-2 BROMIDE COMPLEX
977	1d2v	C	246	732	0	0.63	-1202.08		MYELOPEROXIDASE; CHAIN: A, B; MYELOPEROXIDASE; CHAIN: C, D;	OXIDOREDUCTASE HEME-PROTEIN, PEROXIDASE, OXIDOREDUCTASE, PEROXIDASE-2 BROMIDE COMPLEX
977	1eqg	A	108	710	1.4e-93			148.59	PROSTAGLANDIN H2 SYNTHASE-1; CHAIN: A, B;	OXIDOREDUCTASE COX-1; EGF DOMAIN, NSAID BINDING, IBUPROFEN, MEMBRANE-BINDING 2 DOMAIN, CYCLOOXYGENASE, PEROXIDASE, OXIDOREDUCTASE, 3 DIOXYGENASE, PEROXIDASE
977	1eqg	A	1	584	2.8e-99			149.69	PROSTAGLANDIN H2 SYNTHASE-1; CHAIN: A, B;	OXIDOREDUCTASE COX-1; EGF DOMAIN, NSAID BINDING, IBUPROFEN, MEMBRANE-BINDING 2 DOMAIN, CYCLOOXYGENASE, PEROXIDASE, OXIDOREDUCTASE, 3 DIOXYGENASE, PEROXIDASE
981	1a2y	A	34	146	4.2e-40			59.18	MONOCLONAL ANTIBODY D1.3; CHAIN:	COMPLEX (IMMUNOGLOBULIN/HYDROLAS

626

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A, B; LYSOZYME; CHAIN: C;	B) COMPLEX (IMMUNOGLOBULIN/HYDROLASE, IMMUNOGLOBULIN V2 REGION, SIGNAL, HYDROLASE, GLYCOSIDASE, BACTERIOLYTIC 3 ENZYME, EGG WHITE
981	1a6v	H	34	150	1.1e-17			56.08	B1-8; CHAIN: L, H, M, I, N, J;	IMMUNOGLOBULIN, HAPTEN
981	1a7q	L	1	101	1.4e-31			50.53	MONOCLONAL ANTIBODY D1.3; CHAIN: L, H;	IMMUNOGLOBULIN
981	1a7q	L	34	146	1.4e-38			59.67	MONOCLONAL ANTIBODY D1.3; CHAIN: L, H;	IMMUNOGLOBULIN
981	1ad0	A	34	215	9.8e-60			55.26	FAB FRAGMENT, ANTIBODY A5B7; CHAIN: A, B, C, D;	IMMUNOGLOBULIN
981	1adq	L	36	203	2.8e-68	0.25	-1202.08		IGG4 REA; CHAIN: A; RF-ANIGM/LAMBDA; CHAIN: H, L;	IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
981	1a07	B	36	215	1.5e-12			61.51	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS I MHC, T-CELL RECEPTOR, VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
981	1a9k	L	36	203	1.4e-62	0.10	-1202.08		FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
981	1at1	D	34	146	1.4e-35			56.21	CYTOCHROME C OXIDASE; CHAIN: A, B; ANTIBODY FV FRAGMENT; CHAIN: C, D;	AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN COMPLEX (OXIDOREDUCTASE/ANTIBODY) CYTOCHROME AA3, COMPLEX IV, FERROCYTOCHROME C, COMPLEX (OXIDOREDUCTASE/ANTIBODY), ELECTRON TRANSPORT, 2 TRANSMEMBRANE, CYTOCHROME OXIDASE, ANTIBODY COMPLEX
981	1b0w	A	34	153	2.8e-43			55.87	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C;	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
981	1b2w	L	34	198	2.8e-62	0.16	-1202.08		ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
981	1b6d	A	34	198	4.2e-62	0.10	-1202.08		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
981	1b62	D	35	214	7e-20			55.30	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	
981	1bj1	J	34	198	9.8e-64	0.18	-1202.08		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
981	1bvk	A	34	146	1.4e-44			55.47	HULYS11; CHAIN: A, B, D, E, LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
981	1bw w	A	32	147	5.6e-45			57.52	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENICE-JONES 2 PROTEIN, IMMUNE SYSTEM ANTIBODY THERAPEUTIC, ANTIBODY, CD52
981	1ce1	L	34	215	8.4e-61			55.66	CAMPATH-1H; LIGHT CHAIN; CHAIN: L; CAMPATH-1H; HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	
981	1dee	A	34	198	1.4e-63	0.23	-1202.08		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-1BP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
981	1dftb	L	34	215	5.6e-60			55.95	IMMUNOGLOBULIN 3D6 FAB 1DFB 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
981	1dm0	A	35	198	2.8e-62	0.25	-1202.08		IGM-KAPPA COLD AGGLUTININ (LIGHT CHAIN); CHAIN: A, C; IGM-KAPPA COLD AGGLUTININ (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM FAB, IGM, ANTIBODY, COLD AGGLUTININ, HUMAN
981	1fgv	L	1	102	7e-41			50.49	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4	
981	1fgv	L	34	146	1.3e-46			57.40	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4	
981	1bxc	B	34	215	4.2e-25			57.17	GAMMA-DELTA T-CELL RECEPTOR; CHAIN: A, C, E, G; GAMMA-DELTA T-CELL RECEPTOR; CHAIN: B, D, F, H;	IMMUNE SYSTEM T-CELL RECEPTOR DELTA CHAIN; T-CELL RECEPTOR GAMMA CHAIN; IG DOMAIN, T CELL RECEPTOR, TCR, GPCR
981	1igm	L	34	155	4.2e-46			59.49	IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM3	
981	1jrh	L	34	206	1.4e-45			56.13	ANTIBODY A6; CHAIN: L, H; INTERFERON-GAMMA RECEPTOR ALPHA CHAIN; CHAIN: I;	COMPLEX (ANTIBODY/ANTIGEN) CYTOKINE RECEPTOR, COMPLEX (ANTIBODY/ANTIGEN), 2 TRANSMEMBRANE, GLYCOPROTEIN
981	1hl	A	36	203	8.4e-65	0.07	-1202.08		LAMBDA III BENGE JONES PROTEIN CLE; CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENGE JONES PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
981	1ngp	H	34	215	4.2e-23			56.17	NI G9 (IGG1=LAMBDA=); CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN,
981	1plg	H	34	215	2.8e-18			56.05	IGG2A=KAPPA=, 1PLG 4 CHAIN: L, H; 1PLG 5	IMMUNOGLOBULIN
981	1tcr	A	35	215	7e-20			57.82	ALPHA, BETA T-CELL RECEPTOR CHAIN: A, B;	RECEPTOR TCR, T-CELL, RECEPTOR, TRANSMEMBRANE, GLYCOPROTEIN, SIGNAL
981	2fgw	L	34	198	9.8e-64	0.07	-1202.08		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
981	2fgw	L	34	215	9.8e-64			56.05	IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
981	2pcp	B	34	215	9.8e-19			54.90	IMMUNOGLOBULIN; CHAIN: A, B, C, D;	IMMUNOGLOBULIN IMMUNOGLOBULIN
981	8fab	A	38	203	4.2e-66	0.13	-1202.08		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
981	8fab	A	38	215	4.2e-66			55.65	IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
982	12e8	L	24	224	1.4e-74			96.49	2E8 (IGG1=KAPPA=)	IMMUNOGLOBULIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN
982	1a4j	A	23	238	9.8e-72			98.04	IMMUNOGLOBULIN, DIELS ALDER CATALYTIC ANTIBODY; CHAIN: L, H, A, B;	IMMUNOGLOBULIN ANTIBODY, CATALYTIC ANTIBODY, DIELS ALDER, 2 GERMLINE
982	1a07	D	1	97	1.3e-35			70.96	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS I MHC, T-CELL RECEPTOR, VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
982	1b2w	L	22	224	1.4e-81			98.81	ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
982	1b6d	A	22	222	1.4e-81	0.00	-1202.08		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN KAPPA IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
982	1b6d	A	22	233	1.4e-81			96.63	IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
982	1b88	A	1	97	1.1e-40			83.05	T CELL RECEPTOR V-ALPHA DOMAIN; CHAIN: A, B;	T CELL RECEPTOR TCR; T CELL RECEPTOR, MHC CLASS I, HUMAN IMMUNODEFICIENCY VIRUS, 2 MOLECULAR RECOGNITION

632

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
982	1bd2	D	22	218	3e-74			230.06	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)
982	1bj1	J	22	223	1.1e-83	0.15	-1202.08		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
982	1ce1	L	22	232	1.4e-79			97.53	CAMPATH-1H; LIGHT CHAIN; CHAIN: L; CAMPATH-1H; HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
982	1cic	A	24	224	1.4e-74			96.33	IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
982	1d9k	A	1	97	2.8e-42			79.58	T-CELL RECEPTOR D10 (ALPHA CHAIN); CHAIN: A, E; T-CELL RECEPTOR D10 (BETA CHAIN); CHAIN: B, F; MHC I-AK A CHAIN (ALPHA CHAIN);	IMMUNE SYSTEM MHC I-AK; MHC I-AK; T-CELL RECEPTOR, MHC CLASS II, D10, I-AK



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: C, G; MHC I-AK B CHAIN (BETA CHAIN); CHAIN: D, H; CONALBUMIN PEPTIDE; CHAIN: P, Q;	
982	1dee	A	22	226	1.4e-84			96.48	IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7 Å RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
982	1dhp	A	12	304	5.6e-92			130.53	DIHYDRODIPICOLINATE SYNTHASE; CHAIN: A, B;	SYNTHASE DHDS; SYNTHASE, DIHYDRODIPICOLINATE
982	1dm0	A	22	223	2.8e-79	0.02	-1202.08		IGM-KAPPA COLD AGGLUTININ (LIGHT CHAIN); CHAIN: A, C; IGM-KAPPA COLD AGGLUTININ (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM FAB, IGM, ANTIBODY, COLD AGGLUTININ, HUMAN
982	1etx	L	22	224	1.4e-74			97.96	TAB2; CHAIN: L, M; TAB2; CHAIN: H, I; CYCLIC PEPTIDE; CHAIN: P, Q	COMPLEX (ANTIBODY/ANTIGEN) CROSS-REACTIVITY, PROTEIN-PEPTIDE RECOGNITION
982	1egj	L	24	221	1.4e-73			100.34	CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR; CHAIN: A; ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H	IMMUNE SYSTEM CYTOKINE RECEPTOR COMPLEXED TO AN ANTIBODY
982	1ent	L	22	224	1.3e-75			98.73	IGG ANTIBODY (LIGHT CHAIN); CHAIN: L; IGG ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM ANTI-FULLERENE ANTIBODY, NANOTUBES

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
982	1fkg	L	21	224	7e-75			97.87	IMMUNOGLOBULIN G1 (KAPPA LIGHT CHAIN) FAB' FRAGMENT 1F1G3	
982	1fhs	L	22	224	8.4e-78			96.30	IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
982	1fvd	A	22	223	7e-81	0.31	-1202.08		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
982	1fyt	D	22	218	2.8e-57			204.11	HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR CHAIN: A; HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR-1 CHAIN: B; HEMAGGLUTININ HAI PEPTIDE CHAIN; CHAIN: C; T-CELL RECEPTOR ALPHA CHAIN; CHAIN: D; T-CELL RECEPTOR BETA CHAIN; CHAIN: E;	IMMUNE SYSTEM HLA-DR1, DRA; HLA-DR1, DRB1 0101; TCR HAI.7 ALPHA CHAIN; TCR HAI.7 BETA CHAIN; PROTEIN-PROTEIN COMPLEX, IMMUNOGLOBULIN FOLD
982	1g9m	L	21	227	1.4e-77			97.48	ENVELOPE GLYCOPROTEIN GP120; CHAIN: G; T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: C; ANTIBODY 17B, LIGHT CHAIN;	VIRUS/VIRAL PROTEIN COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB), HIV-1 EXTERIOR 2 ENVELOPE GP120 FROM LABORATORY-ADAPTED ISOLATE, HXBC3, 3 SURFACE T-CELL GLYCOPROTEIN CD4,

635

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PME score	SEQFOL D score	Compound	PDB annotation
982	1i7z	A	24	224	2.8e-76			98.87	CHAIN: L; ANTIBODY 17B, HEAVY CHAIN; CHAIN: H;	ANTIGEN-BINDING FRAGMENT 4 OF HUMAN IMMUNOGLOBULIN 17B
									CHIMERA OF IG KAPPA CHAIN: HUMAN CONSTANT REGION CHAIN: A, C; CHIMERA OF IG GAMMA-1 CHAIN: HUMAN CONSTANT CHAIN: B, D;	IMMUNE SYSTEM IGG FOLD, ANTIBODY, CHIMERA
982	1kb5	A	1	97	5.6e-43			87.06	KB5-C20 T-CELL ANTIGEN RECEPTOR, CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2
										(IMMUNOGLOBULIN/RECEPTOR)
982	1qrm	D	23	221	1.2e-77			239.88	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE P6A; CHAIN: C; T-CELL RECEPTOR, ALPHA CHAIN; CHAIN: D; T-CELL RECEPTOR, BETA CHAIN; CHAIN: E;	IMMUNE SYSTEM MHC CLASS I HLA-A; MHC CLASS I ANTIGEN; A6-TCR; TCR BETA CHAIN; HUMAN TCR/PEPTIDE/MHC COMPLEX, HLA-A2, HTLV-1, TAX, TCR, T 2 CELL RECEPTOR
982	2fgw	L	22	223	7e-83	0.04	-1202.08		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-O2 FAB) 2FGW 4	
983	1b6b	B	221	337	9.8e-10	0.35	-1202.08		ARYLAALKYLAMINE N-	TRANSFERASE

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQPOL D score	Compound	PDB annotation
983	1c1w	A	221	337	1.1e-08	0.14	-1202.08		ACETYLTRANSFERASE; CHAIN: A, B; SEROTONIN N- ACETYLTRANSFERASE; CHAIN: A;	ACETYLTRANSFERASE
983	1ib1	E	221	337	1.1e-08	0.20	-1202.08		14-3-3 ZETA ISOFORM; CHAIN: A, B, C, D; SEROTONIN N- ACETYLTRANSFERASE; CHAIN: E, F, G, H;	SIGNALING PROTEIN/TRANSFERASE PROTEIN KINASE C INHIBITOR PROTEIN-1; ARALKYLAMINE N- ACETYLTRANSFERASE, AA-NAT, N-ACETYL TRANSFERASE, 14-3-3, SIGNAL TRANSDUCTION, PROTEIN-2 PROTEIN COMPLEX, PHOSPHORYLATION
983	1qsm	A	218	334	9.8e-14	0.03	-1202.08		HPA2 HISTONE ACETYLTRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE PROTEIN-ACETYL COENZYME A COMPLEX, ACETYLTRANSFERASE
985	1ext	A	38	209	1.4e-10			53.75	TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALING PROTEIN
985	1ext	A	52	202	1.4e-10	0.25	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALING PROTEIN
985	1g44	A	44	258	3e-07			57.93	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
985	1g44	B	31	250	0.0006			60.13	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
985	1hdj		9	84	1.1e-29			70.79	HUMAN HSP40; CHAIN: NULL;	MOLECULAR CHAPERONE HDJ-1; MOLECULAR CHAPERONE
985	1igr	A	40	171	1.5e-09	0.21	-1202.08		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
985	1klo		44	221	7e-13			55.87	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
985	1ncf	A	52	179	8.4e-09	0.15	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; INCF 4 CHAIN: A, B; INCF 5	SIGNALING PROTEIN TYPE 1 RECEPTOR, STNFR1; INCF 8 BINDING PROTEIN, CYTOKINE INCF 19
985	9wga	A	45	224	7e-14			51.43	LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
987	1a4y	A	127	368	1.2e-40	0.46	-1202.08		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B; E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
987	1a4y	A	246	375	1.2e-19	0.23	-1202.08		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B; E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
987	1a4y	A	54	355	7.5e-45	0.39	-1202.08		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B; E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
987	1a9n	A	108	283	1.3e-28	0.31	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
987	1a9n	A	179	330	1.4e-30	0.79	-1202.08		CHAIN: A, C; U2 B"; CHAIN: B, D;	(NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	A	203	355	1e-32	0.59	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	A	276	368	1.4e-15	0.31	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	A	54	213	1.5e-21	0.15	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	A	83	233	1.2e-29	0.46	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	C	108	283	1.5e-28	0.13	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	C	179	327	9e-31	0.73	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	C	203	355	6e-32	0.50	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
									U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
987	1a9n	C	273	369	1.5e-15	0.36	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	C	54	233	9e-22	0.11	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1a9n	C	83	231	1.5e-29	0.50	-1202.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
987	1d0b	A	110	329	4.5e-37			71.98	INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
987	1d0b	A	170	368	7.5e-34	1.01	-1202.08		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
987	1d0b	A	244	423	1.3e-20	0.78	-1202.08		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
987	1d0b	A	279	433	5.6e-15	0.33	-1202.08		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
987	1d0b	A	33	204	7e-22	0.30	-1202.08		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
987	1d0b	A	48	232	3e-27	0.29	-1202.08		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
987	1d0b	A	61	279	4.5e-37	0.63	-1202.08		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
987	1d0b	A	63	252	2.8e-26	0.41	-1202.08		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
987	1dce	A	272	381	2.8e-13	0.56	-1202.08		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	REPEAT, CALCIUM BINDING, CELL ADHESION TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 Å 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
987	1fo1	B	228	369	9e-17	0.07	-1202.08		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
987	1fqv	A	183	380	9e-20	0.25	-1202.08		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN 2 E3, UBIQUITIN PROTEIN LIGASE
987	1fqv	A	71	400	7.5e-25			69.13	SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN 2 E3, UBIQUITIN PROTEIN LIGASE
987	1fs2	A	187	385	3e-15	0.13	-1202.08		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
987	1fs2	A	53	321	1.5e-18	0.03	-1202.08		SKP2; CHAIN: A, C; SKP1;	LIGASE CYCLIN A/CDK2-



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: B, D;	ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19, SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
987	1yr8	A	126	371	7.5e-28	0.12	-1202.08		GTPASE-ACTIVATING PROTEIN RNAI _SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
								70.31	GTPASE-ACTIVATING PROTEIN RNAI _SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
987	1yr8	A	59	185	3e-14	0.21	-1202.08		GTPASE-ACTIVATING PROTEIN RNAI _SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
987	1yr8	A	75	282	3e-29	0.33	-1202.08		GTPASE-ACTIVATING PROTEIN RNAI _SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-

642

Table 5

SEQ ID No:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
987	1yr6	A	94	330	9e-34	0.48	-1202.08		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
987	2bnh		80	368	1.5e-48	0.52	-1202.08		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
988	1ep2	A	103	359	1.4e-58	0.45	-1202.08		NITROGENASE IRON PROTEIN; CHAIN: A, B;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
988	1ep2	A	104	359	1.4e-58			64.61	NITROGENASE IRON PROTEIN; CHAIN: A, B;	OXIDOREDUCTASE CP2; OXIDOREDUCTASE, NITROGENASE IRON PROTEIN
988	1fp6	A	100	359	9.8e-61			58.49	NITROGENASE IRON PROTEIN; CHAIN: A, B, C, D;	OXIDOREDUCTASE, NITROGENASE IRON PROTEIN
988	1fp6	A	103	359	9.8e-61	0.36	-1202.08		NITROGENASE IRON	OXIDOREDUCTASE, NITROGENASE COMPONENT II, NITROGENASE REDUCTASE, NITROGENASE, NITROGEN FIXATION, NUCLOTIDE, MGADP, FE 2 PROTEIN, AV2

643

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEIN; CHAIN: A, B, C, D;	NITROGENASE COMPONENT II, NITROGENASE REDUCTASE, NITROGENASE, NITROGEN FIXATION, NUCLOTIDE, MGADP, FE 2 PROTEIN, AV2
988	1fp6	A	104	359	3e-46	0.35	-1202.08		NITROGENASE IRON PROTEIN; CHAIN: A, B, C, D;	OXIDOREDUCTASE NITROGENASE COMPONENT II, NITROGENASE REDUCTASE, NITROGENASE, NITROGEN FIXATION, NUCLOTIDE, MGADP, FE 2 PROTEIN, AV2
988	1fts		19	296	1.1e-06			53.22	FTSY; CHAIN: NULL;	SIGNAL RECOGNITION PARTICLE RECEPTOR SIGNAL RECOGNITION PARTICLE RECEPTOR, GTPASE, PROTEIN 2 TARGETING
988	1g20	E	103	359	1.4e-57	0.22	-1202.08		NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN; CHAIN: A, C;	OXIDOREDUCTASE NITROGENASE COMPONENT I, DINITROGENASE, NIFD, NITROGENASE COMPONENT I, DINITROGENASE, NIFK;
									NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN; CHAIN: B, D;	NITROGENASE COMPONENT II, NITROGENASE REDUCTASE, NITROGEN-FIXATION, FE PROTEIN, MOFE PROTEIN, P-CLUSTER AND 2 FEMO COFACTOR
988	1g20	E	104	359	1.4e-57			62.01	NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN; CHAIN: A, C;	OXIDOREDUCTASE NITROGENASE COMPONENT I, DINITROGENASE, NIFD, NITROGENASE COMPONENT I, DINITROGENASE, NIFK;
									NITROGENASE MOLYBDENUM-IRON PROTEIN BETA CHAIN; CHAIN: B, D;	NITROGENASE COMPONENT II, NITROGENASE REDUCTASE, NITROGEN-FIXATION, FE

644

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
988	1g20	E	104	359	4.5e-48	0.58	-1202.08		NITROGENASE IRON PROTEIN; CHAIN: E, F, G, H;	PROTEIN, MORE PROTEIN, P-CLUSTER AND 2 FEMO COFACTOR
									NITROGENASE MOLYBDENUM-IRON PROTEIN ALPHA CHAIN; CHAIN: A, C;	OXIDOREDUCTASE NITROGENASE COMPONENT I, DINITROGENASE, NIFD; NITROGENASE COMPONENT I, DINITROGENASE, NIFK; NITROGENASE COMPONENT II, NITROGENASE REDUCTASE, NITROGEN-FIXATION, FE PROTEIN, MORE PROTEIN, P-CLUSTER AND 2 FEMO COFACTOR
988	1g3q	A	101	350	2.8e-48	0.54	-1202.08		CELL DIVISION INHIBITOR; CHAIN: A;	CELL CYCLE, HYDROLASE MIND ATPASE, ALPHA-BETA-ALPHA LAYERED, PROTEIN-ADP COMPLEX
988	1g3q	A	101	354	2.8e-48			104.71	CELL DIVISION INHIBITOR; CHAIN: A;	CELL CYCLE, HYDROLASE MIND ATPASE, ALPHA-BETA-ALPHA LAYERED, PROTEIN-ADP COMPLEX
988	1hyq	A	102	351	9.8e-46			105.60	CELL DIVISION INHIBITOR (MIND-1); CHAIN: A;	CELL CYCLE MIND; MINC, FTSZ, BACTERIAL CELL DIVISION
988	1hyq	A	103	350	9.8e-46	0.31	-1202.08		CELL DIVISION INHIBITOR (MIND-1); CHAIN: A;	CELL CYCLE MIND; MINC, FTSZ, BACTERIAL CELL DIVISION
988	1hyq	A	104	350	1.5e-39	0.70	-1202.08		CELL DIVISION INHIBITOR (MIND-1); CHAIN: A;	CELL CYCLE MIND; MINC, FTSZ, BACTERIAL CELL DIVISION
988	1j8m	F	26	298	7e-17			50.31	SIGNAL RECOGNITION 54 KDA PROTEIN; CHAIN: F;	SIGNALING PROTEIN SRP54; SIGNALING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
988	In2c	E	100	359	9.8e-61			64.02	NITROGENASE MOLYBDENUM-IRON PROTEIN; CHAIN: A, B, C, D; NITROGENASE IRON PROTEIN; CHAIN: E, F, G, H;	COMPLEX OF NITROGENASE PROTEINS NITROGENASE COMPONENT I, DINITROGENASE; NITROGENASE COMPONENT II, NITROGENASE REDUCTASE; NITROGENASE, NITROGEN FIXATION, SIGNAL TRANSDUCTION, 2 ELECTRON TRANSFER, ATP HYDROLYSIS, COMPLEX OF NITROGENASE 3 PROTEINS
988	In2c	E	103	359	9.8e-61	0.41	-1202.08		NITROGENASE MOLYBDENUM-IRON PROTEIN; CHAIN: A, B, C, D; NITROGENASE IRON PROTEIN; CHAIN: E, F, G, H;	COMPLEX OF NITROGENASE PROTEINS NITROGENASE COMPONENT I, DINITROGENASE; NITROGENASE COMPONENT II, NITROGENASE REDUCTASE; NITROGENASE, NITROGEN FIXATION, SIGNAL TRANSDUCTION, 2 ELECTRON TRANSFER, ATP HYDROLYSIS, COMPLEX OF NITROGENASE 3 PROTEINS
989	1aut	L	605	684	1.4e-11	0.03	-1202.08		ACTIVATED PROTEIN C <sub>3</sub> CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
989	1ceV	A	628	682	1.3e-14	0.42	-1202.08		CHYMOTRYPSIN INHIBITOR; CHAIN: A;	HYDROLASE INHIBITOR AMCI PROTEIN INHIBITOR, HEMOLYMPH, APIS MELLIFERA, CANONICAL 2 INHIBITOR
989	1dva	L	146	229	5.6e-11	0.04	-1202.08		DES-GLA FACTOR VIIA	HYDROLASE/HYDROLASE

646

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE F-76; CHAIN: X, Y;	INHIBITOR PROTEIN-PEPTIDE COMPLEX
989	1fak	L	182	259	4.2e-12	0.02	-1202.08		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX/SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, BGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
989	1hx2	A	628	682	9e-18	0.33	-1202.08		BST1; CHAIN: A;	HYDROLASE INHIBITOR BOMBINA SKIN TRYPSIN INHIBITOR BETA-SHEET DISULFIDE-RICH
989	1qub	A	89	364	6e-15	0.05	-1202.08		HUMAN BETA2-GLYCOPROTEIN I; CHAIN: A;	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
992	1a3r	L	20	221	9.8e-83	0.69	-1202.08		IGG2A; CHAIN: L, H; HUMAN RHINOVIRUS CAPSID PROTEIN VP2; CHAIN: P;	COMPLEX (IMMUNOGLOBULIN/VIRAL PEPTIDE) ANTIBODY 8F5; IMMUNOGLOBULIN, ANTIBODY, RHINOVIRUS, NEUTRALIZATION, 2 CONTINUOUS EPTOPE, COMPLEX (IMMUNOGLOBULIN/VIRAL

647

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
992	1aqk	L	21	221	6e-86	0.91	-1202.08		FAB B7-15A2; CHAIN: L, H <sub>2</sub>	PEPTIDE) IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
992	1aqk	L	21	222	6e-86			232.04	FAB B7-15A2; CHAIN: L, H <sub>2</sub>	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
992	1bjm	A	20	222	3e-83			228.10	LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER; 1BJM 6 CHAIN: A, B; 1BJM 7	IMMUNOGLOBULIN BENGE- JONES PROTEIN; 1BJM 8 BENGE JONES, ANTIBODY, MULTIPLE QUATERNARY STRUCTURES 1BJM 13
992	1bjm	A	21	221	1.3e-74	0.92	-1202.08		LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER; 1BJM 6 CHAIN: A, B; 1BJM 7	IMMUNOGLOBULIN BENGE- JONES PROTEIN; 1BJM 8 BENGE JONES, ANTIBODY, MULTIPLE QUATERNARY STRUCTURES 1BJM 13
992	1bjm	A	21	221	3e-83	1.01	-1202.08		LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER; 1BJM 6 CHAIN: A, B; 1BJM 7	IMMUNOGLOBULIN BENGE- JONES PROTEIN; 1BJM 8 BENGE JONES, ANTIBODY, MULTIPLE QUATERNARY STRUCTURES 1BJM 13
992	1dzb	A	3	129	1.4e-44	0.74	-1202.08		SCFV FRAGMENT 1F9; CHAIN: A, B; TURKEY EGG-WHITE LYSOZYME C; CHAIN: X, Y;	COMPLEX (ANTIBODY ANTIGEN) 1,4-BETA-N- ACETYLMURAMIDASE C; SINGLE-DOMAIN ANTIBODY, TURKEY EGG-WHITE LYSOZYME, 2 ANTIBODY-PROTEIN COMPLEX, SINGLE-CHAIN FV FRAGMENT

648

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
992	1hil	A	20	221	4.2e-83	0.76	-1202.08		IMMUNOGLOBULIN IGG2A FAB FRAGMENT (FAB 17/9) 1HIL 3	
992	1hil	A	20	222	4.2e-83			155.83	IMMUNOGLOBULIN IGG2A FAB FRAGMENT (FAB 17/9) 1HIL 3	
992	1ifh	L	20	221	4.2e-83	0.78	-1202.08		IMMUNOGLOBULIN IGG2A FAB FRAGMENT (FAB 17/9) COMPLEX WITH PEPTIDE OF 1IFH 3 INFLUENZA HEMAGGLUTININ HA1 (STRAIN X47) (RESIDUES 101-107) 1IFH 4	
992	1ifh	L	20	222	4.2e-83			155.98	IMMUNOGLOBULIN IGG2A FAB FRAGMENT (FAB 17/9) COMPLEX WITH PEPTIDE OF 1IFH 3 INFLUENZA HEMAGGLUTININ HA1 (STRAIN X47) (RESIDUES 101-107) 1IFH 4	
992	1nqb	A	12	130	1.4e-45	0.93	-1202.08		SINGLE-CHAIN ANTIBODY FRAGMENT; CHAIN: A, C;	IMMUNOGLOBULIN VARIABLE HEAVY (VH) DOMAIN, VARIABLE LIGHT (VL) ANTIBODY FRAGMENT, MULTIVALENT ANTIBODY, DOMAIN 2 SWAPPING, IMMUNOGLOBULIN
992	1nqk	A	12	129	7e-46	1.09	-1202.08		MFE-23 RECOMBINANT ANTIBODY FRAGMENT; CHAIN: A;	IMMUNOGLOBULIN IMMUNOGLOBULIN, SINGLE-CHAIN FV, ANTI-CARCINOEMBRYONIC 2 ANTIGEN
992	1sbs	L	20	221	1.4e-84	0.73	-1202.08		MONOCLONAL	MONOCLONAL ANTIBODY



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ANTIBODY 3A2; CHAIN: H, L;	MONOCLONAL ANTIBODY, FAB-FRAGMENT, REPRODUCTION
992	1sbs	L	20	222	1.4e-84			158.44	MONOCLONAL ANTIBODY 3A2; CHAIN: H, L;	MONOCLONAL ANTIBODY MONOCLONAL ANTIBODY, FAB-FRAGMENT, REPRODUCTION
992	2fb4	L	20	221	7e-77	0.96	-1202.08		IMMUNOGLOBULIN 2FB4 4	
992	2fb4	L	20	222	7e-77			238.13	IMMUNOGLOBULIN 2FB4 4	
992	2mcg	1	1	98	4.2e-51			88.32	IMMUNOGLOBULIN 2MCG 4	
									LAMBDA LIGHT CHAIN DIMER (/MCGS) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
992	2mcg	1	20	222	1.3e-82			268.98	IMMUNOGLOBULIN 2MCG 4	
									LAMBDA LIGHT CHAIN DIMER (/MCGS) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
992	2mcg	1	21	221	1.3e-82	0.86	-1202.08		IMMUNOGLOBULIN 2MCG 4	
									LAMBDA LIGHT CHAIN DIMER (/MCGS) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
992	7fab	L	20	222	4.5e-78			203.84	IMMUNOGLOBULIN 7FAB 3	
									LAMBDA LIGHT CHAIN DIMER (/MCGS) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
992	7fab	L	21	221	4.5e-78	0.91	-1202.08		IMMUNOGLOBULIN 7FAB 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQIDOL D score	Compound	PDB annotation
994	12e8	L	28	215	4.2e-09			63.88	IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7EAB 3	IMMUNOGLOBULIN
994	1a4j	A	28	232	5.6e-11			63.22	IMMUNOGLOBULIN, DIELS ALDER CATALYTIC ANTIBODY; CHAIN: L, H, A, B;	IMMUNOGLOBULIN ANTIBODY, CATALYTIC ANTIBODY, DIELS ALDER, 2 GERMLINE
994	1ad0	A	28	234	5.6e-13			64.06	FAB FRAGMENT, ANTIBODY A5B7; CHAIN: A, B, C, D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB FRAGMENT
994	1b2w	L	28	234	2.8e-11			68.36	ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
994	1b4j	L	28	232	5.6e-13			68.55	ANTIBODY; CHAIN: L, H;	ANTIBODY ENGINEERING ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODIES, 2 FAB, X-RAY STRUCTURES, GAMMA-INTERFERON
994	1b1h	A	2	306	1.4e-40	0.04	-1202.08		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
994	1cdy		37	213	3e-29	0.19	-1202.08		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
994	1cs6	A	1	308	3e-55			102.00	AXONIN-1; CHAIN: A;	TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
994	1cs6	A	22	304	3e-55	0.30	-1202.08		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
994	1cs6	A	2	307	9.8e-48	0.10	-1202.08		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
994	1cvs	C	127	306	9.8e-46	0.16	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
994	1cvs	C	19	229	4.2e-29			63.89	FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
994	1cvs	C	217	337	7e-23	0.10	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
994	1cvs	D	127	306	9.8e-44	0.13	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR RECEPTOR 1; CHAIN: C, D;	SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
994	1cvs	D	19	213	1.3e-29			73.52	FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
994	1cvs	D	217	337	7e-23	0.18	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
994	1d5i	L	28	231	1.4e-12			63.72	CHIMERIC GERMLINE PRECURSOR OF OXY-COPE CHAIN: L; CHIMERIC GERMLINE PRECURSOR OF OXY-COPE CHAIN: H;	IMMUNE SYSTEM/IMMUNE SYSTEM
994	1dgi	R	8	304	7.5e-48			75.15	POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
994	1eaj	A	36	127	3e-14	0.22	-1202.08		COXSACKIE VIRUS AND ADENOVIRUS RECEPTOR; CHAIN: A, B;	VIRUS/VIRAL PROTEIN RECEPTOR COXSACKIEVIRUS B-ADENOVIRUS RECEPTOR, HCAR, VIRUS/VIRAL PROTEIN

653

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
994	1epf	A	130	298	9e-30	0.21	-1202.08		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	RECEPTOR, IMMUNOGLOBULIN V DOMAIN FOLD, 2 SYMMETRIC DIMER
994	1epf	A	135	290	7e-22	0.19	-1202.08		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
994	1epf	A	29	213	1.1e-34			82.22	NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
994	1epf	A	37	213	1.1e-34	0.28	-1202.08		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
994	1ev2	E	20	213	6e-29			71.72	FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
994	1ev2	G	128	310	4.2e-44	0.22	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
994	1ev2	G	20	217	3e-30			70.77	FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
994	1ev2	G	30	217	3e-30	0.02	-1202.08		FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH

654

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
994	1evt	C	19	213	9.8e-30			76.08	FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
994	1f2q	A	129	307	1.5e-32	0.18	-1202.08		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
994	1f6a	A	125	307	1.5e-37	0.19	-1202.08		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
994	1f97	A	128	315	9e-37			78.34	JUNCTION ADHESION MOLECULE; CHAIN: A;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
994	1f97	A	129	304	9e-37	0.17	-1202.08		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
994	1f97	A	136	301	1.4e-28	0.19	-1202.08		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD

655

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
994	1f97	A	31	203	8.4e-29	0.05	-1202.08		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
994	1fcg	A	128	304	6e-35	0.07	-1202.08		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
994	1fcg	A	25	215	1.5e-24			64.84	FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
994	1fcg	A	37	214	1.5e-24	0.13	-1202.08		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
994	1fhl	A	125	309	3e-36	0.11	-1202.08		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
994	1fsk	B	28	215	1.1e-10			64.25	MAJOR POLLEN ALLERGEN BET V 1-A; CHAIN: A, D, G, J; IMMUNOGLOBULIN KAPPA LIGHT CHAIN; CHAIN: B, E, H, K; ANTIBODY HEAVY CHAIN FAB; CHAIN: C, F, L, L;	IMMUNE SYSTEM BET V 1-A, BETV1 ALLERGEN, BV16 FAB- FRAGMENT, KAPPA MOPC21 CODING SEQUENCE, HEAVY CHAIN OF THE MONOCLONAL ANTIBODY MST2; BET V 1, BV16 FAB FRAGMENT, ANTIBODY ALLERGEN COMPLEX
994	1g9m	L	28	232	1.4e-11			64.69	ENVELOPE GLYCOPROTEIN GP120; CHAIN: G; T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: C; ANTIBODY	VIRUS/VIRAL PROTEIN COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB), HIV-1 EXTERIOR 2 ENVELOPE GP120 FROM LABORATORY-ADAPTED ISOLATE, HXBC2, 3 SURFACE T-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									17B, LIGHT CHAIN; CHAIN: L; ANTIBODY 17B, HEAVY CHAIN; CHAIN: H; MUTANT AL2 6E7S9G; CHAIN: A;	CELL GLYCOPROTEIN CD4, ANTIGEN-BINDING FRAGMENT 4 OF HUMAN IMMUNOGLOBULIN 17B
994	1h8n	A	27	245	1.4e-10			64.09	NEURAL CELL ADHESION MOLECULE; CHAIN: A;	CELL ADHESION N-CAM; INTERMEDIATE IMMUNOGLOBULIN FOLD FRAMEWORK
994	lie5	A	19	128	2.8e-15	0.15	-1202.08		NEURAL CELL ADHESION MOLECULE; CHAIN: A;	CELL ADHESION N-CAM; INTERMEDIATE IMMUNOGLOBULIN FOLD
994	lie5	A	207	304	1.3e-22	0.22	-1202.08		NEURAL CELL ADHESION MOLECULE; CHAIN: A;	CELL ADHESION N-CAM; INTERMEDIATE IMMUNOGLOBULIN FOLD
994	liil	G	128	310	1.1e-44	0.08	-1202.08		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF- 2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
994	liil	G	30	217	7.5e-28	0.07	-1202.08		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF- 2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
994	liib	B	30	309	1.5e-48			72.20	INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
994	lkoa		26	126	4.2e-12	0.09	-1202.08		TWITCHIN; CHAIN:	KINASE KINASE, TWITCHIN,



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
994	Inct		30	126	1.3e-14	0.03	-1202.08		NULL; TTIN; CHAIN: NULL;	INTRASTERIC REGULATION MUSCLE PROTEIN CONNECTIN, NEXTMS; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL. 3 MUSCLE PROTEIN
994	Inkr		129	301	3e-27	0.29	-1202.08		P58-CT42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
994	Itmm		30	126	1.3e-14	0.19	-1202.08		MUSCLE PROTEIN TTIN MODULE M5 (CONNECTIN) ITNM 3 (NMR, MINIMIZED AVERAGE STRUCTURE)	
994	2dli	A	128	301	1.5e-28	0.06	-1202.08		ITNM 4 ITNM 58	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
994	2fcb	A	128	307	3e-35	0.10	-1202.08		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
994	2fcb	A	25	217	1.5e-25			67.36	FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
994	3fct	A	28	233	4.2e-12			64.42	METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: A, C; METAL	IMMUNE SYSTEM METAL CHELATASE, CATALYTIC ANTIBODY, FAB FRAGMENT,

658

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
994	3hfm	H	28	215	4.2e-10			65.01	CHELATASE CATALYTIC ANTIBODY; CHAIN: B, D; COMPLEX(ANTIBODY-ANTIGEN) I G*G1 FAB FRAGMENT (HY/HEL\$-10) AND LYSOZYME (E.C.3.2.1.17) 3HFM 4 COMPLEX 3HFM 5	IMMUNE 2 SYSTEM
995	1c1g	A	1	245	1.3e-25			61.11	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
995	1cfe		30	177	2.8e-35	0.52	-1202.08		PATHOGENESIS-RELATED PROTEIN P14A; CHAIN: NULL;	PATHOGENESIS-RELATED PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
995	1cfe		30	177	7.5e-40			91.26	PATHOGENESIS-RELATED PROTEIN P14A; CHAIN: NULL;	PATHOGENESIS-RELATED PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
995	1cfe		32	177	7.5e-40	0.74	-1202.08		PATHOGENESIS-RELATED PROTEIN P14A; CHAIN: NULL;	PATHOGENESIS-RELATED PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
995	1ij2	J	1	132	2.8e-46			114.02	23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A;	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									RIBOSOMAL PROTEIN L3; CHAIN: B;	PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P,
									RIBOSOMAL PROTEIN L4; CHAIN: C;	HMAL5, HL13; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10; 50S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN P0,
									RIBOSOMAL PROTEIN L5; CHAIN: D;	HMAL10, L10E; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P,
									RIBOSOMAL PROTEIN L6; CHAIN: E;	HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P,
									RIBOSOMAL PROTEIN L7AE; CHAIN: F;	HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E,
									RIBOSOMAL PROTEIN L10; CHAIN: G;	HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P,
									RIBOSOMAL PROTEIN L10E; CHAIN: H;	HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E,
									RIBOSOMAL PROTEIN L13; CHAIN: I;	HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P,
									RIBOSOMAL PROTEIN L14; CHAIN: J;	HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA RIBOSOME ASSEMBLY, RNA-
									RIBOSOMAL PROTEIN L15; CHAIN: K;	
									RIBOSOMAL PROTEIN L15E; CHAIN: L;	
									RIBOSOMAL PROTEIN L18; CHAIN: M;	
									RIBOSOMAL PROTEIN L18E; CHAIN: N;	
									RIBOSOMAL PROTEIN L19E; CHAIN: O;	
									RIBOSOMAL PROTEIN L21E; CHAIN: P;	
									RIBOSOMAL PROTEIN L22; CHAIN: Q;	
									RIBOSOMAL PROTEIN L23; CHAIN: R;	
									RIBOSOMAL PROTEIN	

660

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									L24; CHAIN: S; RIBOSOMAL PROTEIN L24E; CHAIN: T; RIBOSOMAL PROTEIN L29; CHAIN: U; RIBOSOMAL PROTEIN L30; CHAIN: V; RIBOSOMAL PROTEIN L31E; CHAIN: W; RIBOSOMAL PROTEIN L32E; CHAIN: X; RIBOSOMAL PROTEIN L37AE; CHAIN: Y; RIBOSOMAL PROTEIN L37E; CHAIN: Z; RIBOSOMAL PROTEIN L39E; CHAIN: 1; RIBOSOMAL PROTEIN L44E; CHAIN: 2;	RNA, PROTEIN-RNA, PROTEIN- PROTEIN
995	1qnx	A	1	177	1.3e-39	0.37	-1202.08		VES V 5; CHAIN: A;	ALLERGEN ANTIGEN 5; ANTIGEN 5, ALLERGEN, VESPID VENOM
995	1whi		12	132	4.2e-52			71.98	RIBOSOMAL PROTEIN L14; CHAIN: NULL;	RIBOSOMAL PROTEIN RIBOSOMAL PROTEIN, RRNA- BINDING
996	1cs6	A	55	417	2.8e-46			79.13	AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
996	1f2q	A	58	225	9.8e-36			140.65	HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RL- ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
996	1f6a	A	56	225	4.2e-36			141.95	HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IGE EPSILON	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN C REGION; CHAIN: B, D;	RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
996	1fcg	A	55	224	1.4e-36			135.32	FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
996	1fhl	A	54	229	7e-34			127.96	LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
996	1qnx	A	1	180	9.8e-46			89.64	VES V 5; CHAIN: A;	ALLERGEN ANTIGEN 5; ANTIGEN 5, ALLERGEN, VESPID VENOM
996	2fcb	A	55	226	2.8e-38			135.92	FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
999	1byg	A	1	271	1.3e-59			81.76	C-TERMINAL SRC KINASE; CHAIN: A;	TRANSFERASE CSK; PROTEIN KINASE, C-TERMINAL SRC KINASE, PHOSPHORYLATION, 2
999	1e3h	A	2	476	1.4e-87			148.05	GUANOSINE PENTAPHOSPHATE SYNTHETASE; CHAIN: A;	STAUROSPORINE, TRANSFERASE POLYRIBONUCLEOTIDE TRANSFERASE POLYNUCLEOTIDE PHOSPHORYLASE, GUANOSINE POLYRIBONUCLEOTIDE TRANSFERASE, ATP-GTP 2
999	1e3h	A	9	474	1.4e-87	0.11	-1202.08		GUANOSINE PENTAPHOSPHATE SYNTHETASE; CHAIN: A;	DIPHOSPHOTRANSFERASE, RNA PROCESSING, RNA DEGRADATION POLYRIBONUCLEOTIDE TRANSFERASE POLYNUCLEOTIDE PHOSPHORYLASE, GUANOSINE POLYRIBONUCLEOTIDE TRANSFERASE, ATP-GTP 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										DIPHOSPHOTRANSFERASE, RNA PROCESSING, RNA DEGRADATION
999	1e3p	A	2	476	1.4e-87			140.99	GUANOSINE PENTAPHOSPHATE SYNTHETASE; CHAIN: A;	POLYRIBONUCLEOTIDE TRANSFERASE POLYNUCLEOTIDE PHOSPHORYLASE, GUANOSINE POLYRIBONUCLEOTIDE TRANSFERASE, ATP-GTP DIPHOSPHOTRANSFERASE, 2 RNA PROCESSING, RNA DEGRADATION
999	1e3p	A	9	474	1.4e-87	0.17	-1202.08		GUANOSINE PENTAPHOSPHATE SYNTHETASE; CHAIN: A;	POLYRIBONUCLEOTIDE TRANSFERASE POLYNUCLEOTIDE PHOSPHORYLASE, GUANOSINE POLYRIBONUCLEOTIDE TRANSFERASE, ATP-GTP DIPHOSPHOTRANSFERASE, 2 RNA PROCESSING, RNA DEGRADATION
999	1etj	A	12	344	2.8e-39			82.19	TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
999	1fgk	A	4	271	1.3e-55			101.99	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
999	1fgk	B	2	271	2.8e-56			99.67	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
999	1got	B	5	343	2.8e-51			75.43	GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
999	1iep	A	1	271	5.6e-74			115.77	PROTO-ONCOGENE TYROSINE-PROTEIN KINASE ABL; CHAIN: A, B;	TRANSFERASE P150, C-ABL; KINASE, KINASE INHIBITOR, STI-571, ACTIVATION LOOP
999	1it3	A	1	271	1.3e-55			90.03	INSULIN RECEPTOR; CHAIN: A; PEPTIDE SUBSTRATE; CHAIN: B;	COMPLEX (TRANSFERASE/SUBSTRATE) TYROSINE KINASE, SIGNAL TRANSDUCTION, PHOSPHOTRANSFERASE, 2 COMPLEX (KINASE/PEPTIDE SUBSTRATE/ATP ANALOG), ENZYME, 3 COMPLEX (TRANSFERASE/SUBSTRATE)
999	1qpc	A	3	270	9.8e-77			105.38	LCK KINASE; CHAIN: A;	TRANSFERASE ALPHA BETA FOLD
1000	1aln	A	17	117	9.8e-45	0.69		-1202.08	B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	COMPLEX (ANTIGEN/PEPTIDE) B35, MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1000	1aca	B	22	115	2.8e-39	0.54		-1202.08	HLA-DR3; CHAIN: A, B; CLIP; CHAIN: C;	COMPLEX (TRANSMEMBRANE/GLYCOPROTEIN) MHC GLYCOPROTEIN,

664

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQEQL D score	Compound	PDB annotation
1000	1agd	A	17	117	4.2e-44	0.36	-1202.08		B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKKYL - INDEX PEPTIDE); CHAIN: C;	COMPLEX (TRANSMEMBRANE/GLYCOPROTEIN)
1000	1aqd	B	22	115	2.8e-39	0.78	-1202.08		HLA-DRI CLASS II HISTOCOMPATIBILITY PROTEIN; CHAIN: A, B, D, E, G, H, J, K; HLA-A2; CHAIN: C, F, I, L;	COMPLEX (MHC PROTEIN/ANTIGEN) DRA, DRB1 01010; COMPLEX (MHC PROTEIN/ANTIGEN); HISTOCOMPATIBILITY ANTIGEN
1000	1efx	A	17	118	1.4e-44	0.54	-1202.08		HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1000	1efk	B	1	85	7e-32			129.28	H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, K-B CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; NUCLEOCAPSID PROTEIN; CHAIN: P;	IMMUNE SYSTEM SEV9; MAJOR HISTOCOMPATIBILITY COMPLEX PEPTIDE-MHC
1000	1efk	B	21	119	1.4e-35			166.40	H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, K-B CHAIN:	IMMUNE SYSTEM SEV9; MAJOR HISTOCOMPATIBILITY COMPLEX PEPTIDE-MHC



Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A; BETA-2-MICROGLOBULIN; CHAIN: B; NUCLEOCAPSID PROTEIN; CHAIN: P;	
1000	1hsa	A	17	117	9.8e-45	0.39	-1202.08		HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN 1HSA 3 /HLA-B(ASTERISK)2705\$ 1HSA 4	
1000	1i4f	A	17	116	8.4e-42	0.46	-1202.08		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM MAGP-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1000	1jk8	B	22	115	2.8e-40	0.24	-1202.08		MHC CLASS II HLA-DQ8; CHAIN: A; MHC CLASS II HLA-DQ8; CHAIN: B; INSULIN B PEPTIDE; CHAIN: C;	IMMUNE SYSTEM HLA-DQ8, INSULIN B PEPTIDE, TYPE 1 DIABETES, AUTOIMMUNITY
1000	1mh6	A	17	115	8.4e-40	0.60	-1202.08		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-E; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D; PEPTIDE (VMARPTVLL); CHAIN: P, Q;	MAJOR HISTOCOMPATIBILITY COMPLEX MHC NONCLASSICAL CHAIN, MHC-E, HLA-E, MHC CLASS HLA-E, HLA E, MAJOR HISTOCOMPATIBILITY COMPLEX, MHC, HLA, 2 BETA 2 MICROGLOBULIN, PEPTIDE, LEADER PEPTIDE, 3 NON-CLASSICAL MHC, CLASS IB MHC
1000	1qgd	A	17	115	1.3e-42	0.65	-1202.08		HISTOCOMPATIBILITY	IMMUNE SYSTEM

666

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1000	3ftu	B	1	85	5.6e-32			120.95	NEONATAL FC RECEPTOR; CHAIN: A, C, E; BETA-2-MICROGLOBULIN; CHAIN: B, D, F;	FCRN, BRAMBELL RECEPTOR; COMPLEX (IMMUNOGLOBULIN/BINDING PROTEIN)
1000	3ftu	B	21	119	1.4e-35			157.46	NEONATAL FC RECEPTOR; CHAIN: A, C, E; BETA-2-MICROGLOBULIN; CHAIN: B, D, F;	FCRN, BRAMBELL RECEPTOR; COMPLEX (IMMUNOGLOBULIN/BINDING PROTEIN)
1007	1a7i		390	443	5.1e-15	0.43	0.58		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1007	1a7i		390	447	5.6e-10	0.17	0.66		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1007	1a7i		448	507	5.6e-14	-0.10	0.57		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1007	1a7i		449	506	1.7e-16	0.08	0.58		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1007	1a7i		510	575	3.4e-12	-0.05	0.30		QCRP2 (LIM1); CHAIN: NULL;	BINDING PROTEIN, ZINC 2 FINGER
1007	1a7i		510	577	9.8e-10	0.38	0.76		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1007	1a7i		510	577	9.8e-10	0.38	0.76		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1007	1b8t	A	383	572	1.2e-33	-0.26	0.03		CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
1007	1b8t	A	449	577	8.4e-26	0.06	0.07		CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
1007	1c1l		382	443	1.7e-16	-0.38	0.33		AVIAN CYSTEINE RICH PROTEIN; ICTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS
1007	1c1l		382	449	2.8e-13	-0.54	0.18		AVIAN CYSTEINE RICH PROTEIN; ICTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS
1007	1c1l		446	502	6.8e-14	-0.12	0.51		AVIAN CYSTEINE RICH PROTEIN; ICTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS
1007	1c1l		450	516	7e-16	-0.26	0.05		AVIAN CYSTEINE RICH PROTEIN; ICTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS
1007	1c1l		508	572	1.5e-13	-0.11	0.48		AVIAN CYSTEINE RICH PROTEIN; ICTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1007	1ct1		510	577	2.8e-12	0.16	0.55		AVIAN CYSTEINE RICH PROTEIN, ICTL 3	ICTL 15 METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS
1007	1cxx	A	388	443	8.5e-16	0.52	0.68		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2, CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1007	1cxx	A	388	446	4.2e-12	0.26	0.19		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2, CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1007	1cxx	A	448	504	1.4e-14	0.45	0.78		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2, CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1007	1cxx	A	507	572	8.5e-14	-0.05	0.25		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2, CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1007	1cxx	A	507	574	9.8e-12	-0.37	0.89		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2, CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1007	1g47	A	381	450	2.8e-14	0.05	0.00		PINCH PROTEIN; CHAIN: A;	CELL ADHESION PARTICULARLY INTERESTING NEW CYS-HIS PROTEIN; LIM DOMAIN; ZN FINGER
1007	1g47	A	510	577	1.4e-07	0.53	-0.09		PINCH PROTEIN; CHAIN: A;	CELL ADHESION PARTICULARLY INTERESTING NEW CYS-HIS PROTEIN; LIM DOMAIN; ZN FINGER
1007	1iml		388	449	5.6e-12	0.37	0.57		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NUL1;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
1007	1iml		390	461	3.4e-16	-0.02	0.57		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NUL1;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
1007	1iml		449	512	1.4e-14	0.24	0.00		CYSTEINE RICH	METAL-BINDING PROTEIN CRIP;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
1007	1iml		449	513	5.1e-16	0.01	0.47		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP, METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
1007	1iml		508	577	5.6e-12	0.07	1.00		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP, METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
1007	1zfo		508	541	0.00056	-0.34	0.15		LASP-1; CHAIN: NULL;	METAL-BINDING PROTEIN LIM DOMAIN, ZINC-FINGER, METAL- BINDING PROTEIN
1012	1a88	A	2	149	4.2e-27	0.54	0.00		CHLOROPEPOXIDASE L; CHAIN: A, B, C;	HALOPEPOXIDASE BROMOPEPOXIDASE L, HALOPEPOXIDASE L; HALOPEPOXIDASE, OXIDOREDUCTASE
1012	1a8q		2	156	4.2e-24	0.36	-0.07		BROMOPEPOXIDASE AI; CHAIN: NULL;	HALOPEPOXIDASE CHLOROPEPOXIDASE A1, HALOPEPOXIDASE A1; HALOPEPOXIDASE, OXIDOREDUCTASE
1012	1a8s		2	149	1.4e-26	0.53	0.39		CHLOROPEPOXIDASE F; CHAIN: NULL;	HALOPEPOXIDASE HALOPEPOXIDASE F; HALOPEPOXIDASE, OXIDOREDUCTASE, PROPIONATE COMPLEX
1012	1auo	A	13	213	1.4e-24	0.45	1.00		CARBOXYLESTERASE; CHAIN: A, B;	HYDROLASE HYDROLASE
1012	1auo	A	13	229	1.8e-43	0.70	1.00		CARBOXYLESTERASE; CHAIN: A, B;	HYDROLASE HYDROLASE
1012	1auo	A	13	229	5.1e-31	0.53	1.00		CARBOXYLESTERASE; CHAIN: A, B;	HYDROLASE HYDROLASE
1012	1auo	A	8	214	1.4e-24			106.42	CARBOXYLESTERASE; CHAIN: A, B;	HYDROLASE HYDROLASE

670

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1012	1aao	A	8	230	1.8e-43			123.46	CARBOXYLESTERASE; CHAIN: A, B;	HYDROLASE HYDROLASE
1012	1b6g		6	192	7e-19	0.02	0.09		HALOALKANE DEHALOGENASE; CHAIN: NULL;	HYDROLASE HYDROLASE, HALOALKANE DEHALOGENASE, ALPHA/BETA-HYDROLASE
1012	1bn6	A	6	146	5.6e-22	0.18	0.62		HALOALKANE DEHALOGENASE; CHAIN: A;	HYDROLASE DEHALOGENASE, ALPHA/BETA-HYDROLASE, DHLA, CRYSTAL STRUCTURE
1012	1c4x	A	9	157	8.4e-21	0.44	0.62		2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE CHAIN: A;	HYDROLASE BPHD; HYDROLASE, PCB DEGRADATION
1012	1c7j	A	7	198	3.4e-31	0.24	-0.07		PARA-NITROBENZYL ESTERASE; CHAIN: A;	HYDROLASE PNB ESTERASE; ALPHA-BETA HYDROLASE, DIRECTED EVOLUTION, ORGANIC ACTIVITY, 2 PNB ESTERASE
1012	1cv2	A	12	144	2.8e-15	0.24	-0.05		HALOALKANE DEHALOGENASE; CHAIN: A;	HYDROLASE LINB, 1,3,4,6-TETRACHLORO-1,4-CYCLOHEXADIENE DEHALOGENASE, LINDANE, BIODEGRADATION, ALPHA/BETA-HYDROLASE
1012	1cv1		18	128	5.6e-09	0.60	0.19		TRIACYLGLYCEROL HYDROLASE; CHAIN: NULL;	HYDROLASE TRIACYLGLYCEROL-HYDROLASE, X-RAY CRYSTALLOGRAPHY, 2 PSEUDOMONADACEAE, OXYANION, CIS-PEPTIDE, HYDROLASE
1012	1din		5	199	7e-27	0.59	0.94		DIENELACTONE HYDROLASE; CHAIN: NULL;	HYDROLYTIC ENZYME DLH; DIENELACTONE HYDROLASE, AROMATIC HYDROCARBON CATABOLISM, 2 SERINE ESTERASE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1012	1dqr	A	96	215	0.00034	0.31	0.71		ANTIGEN 85-C; CHAIN: A, B;	CARBOXYMETHYLENEBUTENOLIDE, 3 HYDROLYTIC ENZYME IMMUNE SYSTEM 85C ANTIGEN, 85C, MYCOBACTERIUM TUBERCULOSIS, FIBRONECTIN
1012	1ea5	A	3	223	1.7e-36	0.22	0.37		ACETYLCHOLINESTERASE; CHAIN: A;	CHOLINESTERASE SERINE HYDROLASE, NEUROTRANSMITTER CLEAVAGE, CATALYTIC 2 TRIAD, ALPHA/BETA HYDROLASE
1012	1ehy	A	13	172	1.4e-20	0.53	0.16		SOLUBLE EPOXIDE HYDROLASE; CHAIN: A, B, C, D;	HYDROLASE HYDROLASE, ALPHA/BETA HYDROLASE FOLD, EPOXIDE DEGRADATION, 2 EPICHLOROHYDRIN
1012	1ek1	A	2	149	2.8e-25	0.51	0.48		EPOXIDE HYDROLASE; CHAIN: A, B;	HYDROLASE HOMODIMER, ALPHA/BETA HYDROLASE FOLD, DISUBSTITUTED UREA 2 INHIBITOR
1012	1ekl	B	2	149	2.8e-25	0.64	0.55		EPOXIDE HYDROLASE; CHAIN: A, B;	HYDROLASE HOMODIMER, ALPHA/BETA HYDROLASE FOLD, DISUBSTITUTED UREA 2 INHIBITOR
1012	1eth	A	25	167	0.00014	0.52	0.03		TRIACYLGLYCEROL ACYL-HYDROLASE; CHAIN: A, C; COLIPASE; CHAIN: B, D	COMPLEX (HYDROLASE/COFACTOR) TRIACYLGLYCEROL LIPASE; COMPLEX (HYDROLASE/COFACTOR), LIPID DEGRADATION
1012	1f6w	A	2	222	5.1e-34	0.33	-0.08		BILE SALT ACTIVATED LIPASE; CHAIN: A;	HYDROLASE BILE SALT ACTIVATED LIPASE, ESTERASE, CATALYTIC DOMAIN
1012	1fz2	A	13	214	7e-32	0.84	1.00		ACYL PROTEIN THIOESTERASE 1; CHAIN: A, B;	HYDROLASE ALPHA/BETA HYDROLASE, SERINE HYDROLASE, SAD, ANOMALOUS

672

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
1012	1fj2	A	13	231	3.4e-38	0.52	1.00		ACYL PROTEIN THIOESTERASE 1; CHAIN: A, B;	2 DIFFRACTION HYDROLASE ALPHA/BETA HYDROLASE, SERINE HYDROLASE, SAD, ANOMALOUS 2 DIFFRACTION
1012	1fj2	A	1	215	7e-32			116.80	ACYL PROTEIN THIOESTERASE 1; CHAIN: A, B;	HYDROLASE ALPHA/BETA HYDROLASE, SERINE HYDROLASE, SAD, ANOMALOUS 2 DIFFRACTION
1012	1gpl		25	167	0.00014	0.39	0.11		RP2 LIPASE; CHAIN: NULL;	SERINE ESTERASE RELATED PROTEIN 2 LIPASE, SERINE ESTERASE, HYDROLASE, LIPID DEGRADATION, PANCREAS, 2 GLYCOPROTEIN, CHIMERIC
1012	1i6w	A	23	166	8.4e-09	0.19	0.21		LIPASE A; CHAIN: A, B;	HYDROLASE ALPHA/BETA HYDROLASE
1012	1jkm	A	5	157	7e-12	0.39	0.00		BREFELDIN A ESTERASE; CHAIN: A, B;	SERINE HYDROLASE SERINE HYDROLASE, DEGRADATION OF BREFELDIN A, ALPHA/BETA 2 HYDROLASE FAMILY
1012	1lpb	B	25	167	5.6e-05	0.25	0.04		HYDROLASE/CARBOXY LIC ESTERASE) LIPASE (E.C.3.1.1.3) COMPLEXED WITH COLIPASE AND INHIBITED 1LPB 3 BY UNDECANE PHOSPHONATE METHYL ESTER (TWO CONFORMATIONS) 1LPB 4	
1012	1maa	A	3	221	5.1e-36	0.30	0.46		ACETYLCHOLINESTERASE; CHAIN: A, B, C, D;	HYDROLASE MACHE: HYDROLASE, SERINE ESTERASE, ACETYLCHOLINESTERASE, TETRAMER, 2 HYDROLASE FOLD,



673

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1012	1qe3	A	7	198	3.4e-31	0.09	-0.07		PARA-NITROBENZYL ESTERASE; CHAIN: A;	GLYCOSYLATED PROTEIN
1012	1qge	D	18	128	5.6e-09	0.46	0.30		TRIACYLGLYCEROL HYDROLASE; CHAIN: D; TRIACYLGLYCEROL HYDROLASE; CHAIN: E;	HYDROLASE PSEUDOMONADACEAE, CIS-PEPTIDE, CLOSED CONFORMATION, 2 HYDROLASE, LID
1012	1qi4	A	25	159	5.6e-09	0.23	-0.05		HYDROXYNITRILE LYASE; CHAIN: A;	LYASE OXYNITRILE LYASE; OXYNITRILE, CYANOGENESIS, CYANHYDRIN FORMATION, LYASE
1012	1qtr	A	1	130	2.8e-13	0.32	-0.02		PROLYL AMINOPEPTIDASE; CHAIN: A;	HYDROLASE ALPHA BETA HYDROLASE FOLD, PROLINE, PROLYL AMINOPEPTIDASE, 2 SERRATIA, IMINOPEPTIDASE
1012	1thg		3	204	8.5e-30	0.05	0.07		HYDROLASE(CARBOXY LIC ESTERASE) LIPASE (E.C.3.1.1.3) TRIACYLGLYCEROL HYDROLASE ITHG3	
1012	2bce		2	216	3.4e-35	0.32	-0.08		CHOLESTEROL ESTERASE; CHAIN: NULL;	HYDROLASE BILE SALT ACTIVATED LIPASE, BILE SALT STIMULATED HYDROLASE, SERINE ESTERASE, LIPASE
1012	4lip	D	18	128	1.4e-09	0.35	0.13		TRIACYL-GLYCEROL-HYDROLASE; CHAIN: D, E;	LIPASE LIPASE; LIPASE, HYDROLASE, PSEUDOMONADACEAE, COVALENT INTERMEDIATE, 2 TRIGLYCERIDE ANALOGUE, ENANTIOSELECTIVITY
1017	1a1h	A	211	293	1.2e-24	-0.05	1.00		OGSR ZINC FINGER PEPTIDE; CHAIN: A;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA),

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	ZINC FINGER, DNA-BINDING PROTEIN
1017	1alh	A	239	319	2.8e-29	-0.41	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
									QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1017	1alh	A	239	348	5.1e-23	-0.37	0.51		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
									QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1017	1alh	A	463	543	3.4e-42	0.10	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
									QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1017	1alh	A	58	140	1.4e-25			77.51	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
									QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1017	1alh	A	72	154	1.4e-27			54.94	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
									QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1017	1alh	A	87	154	2.8e-31			57.17	QGSR ZINC FINGER	COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PEPTIDE; CHAIN: A; DUPEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1017	1alh	A	90	169	1.1e-23	-0.34	0.00		QGR ZINC FINGER PEPTIDE; CHAIN: A; DUPEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1017	Imey	C	116	198	2.8e-39	-0.24	0.17		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	172	263	7e-43	-0.07	0.92		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	210	291	9.8e-47	-0.07	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	238	319	2.8e-47	-0.33	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	266	347	8.4e-48	-0.03	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE,

676

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1017	Imey	C	294	375	1.3e-48	0.16	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	29	111	2.8e-50			98.30	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	322	403	8.4e-50	0.63	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	350	431	2.8e-50	0.22	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	378	459	7e-51	0.23	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	406	487	5.6e-50	0.27	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	434	515	1.1e-49	0.09	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN,

677

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: C, F, G;	2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	462	543	4.2e-50	0.16	1.00		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	462	544	4.2e-50			102.09	DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	490	565	1.4e-43	0.07	1.00		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	57	139	2.8e-50			100.99	DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	71	153	2.8e-50			69.61	DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	86	154	4.2e-50			76.26	DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	Imey	C	88	169	1.4e-37	-0.48	0.15		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FINGER PROTEIN; CHAIN: C, F, G;	INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	1mev	G	114	141	1.4e-09	-0.61	0.01		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	1mev	G	208	235	2.8e-11	0.29	0.63		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	1mev	G	264	291	4.2e-12	-0.13	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1017	1tf3	A	173	259	4.2e-17	-0.08	0.11		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
1017	1tf6	A	147	403	1.4e-49	-0.57	0.03		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tf6	A	173	328	2.8e-34	-0.35	0.04		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX

679

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQ/OL D score	Compound	PDB annotation
									CHAIN: B, C, E, F;	(TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	1	153	1.4e-36			90.49	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	211	431	3.4e-55	-0.47	0.21		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	267	412	4.2e-38	0.01	0.90		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	2	153	4.2e-37			91.96	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	323	487	3.4e-68	-0.14	0.76		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: B, C, E, F;	(TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	350	520	6.8e-72			108.86	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	351	496	2.8e-39	0.13	1.00		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	379	543	6.8e-72	-0.19	1.00		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	407	551	2.8e-38	-0.00	0.99		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	435	565	1.4e-32	0.08	1.00		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX



681

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: B, C, E, F;	(TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	7	154	8.4e-37			68.08	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1tt6	A	89	249	8.4e-31	0.13	0.24		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1017	1ubd	C	124	235	1.4e-27	-0.45	0.07		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG I; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	180	291	2.8e-30	-0.19	0.80		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG I; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1017	1ubd	C	1	111	2.8e-35			89.38	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	237	375	3.4e-36	-0.42	0.81		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	246	347	9.8e-34	-0.24	0.90		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	265	403	3.4e-47	-0.33	0.62		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	33	140	4.2e-36			66.22	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DNA; CHAIN: A, B;	INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	348	460	6.8e-53	-0.08	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	3	111	1.4e-36			82.90	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	433	544	1.7e-51	0.11	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	439	543	4.2e-35	0.01	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1017	1ubd	C	467	563	3.4e-45	0.02	0.83		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	(TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	1ubd	C	470	566	2.8e-32	0.17	0.96		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1017	2adr		177	238	3.4e-12	-0.04	0.19		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
1017	2gli	A	144	293	5.6e-30	-0.33	0.24		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	177	405	1.5e-51	-0.61	0.43		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	1	139	9.8e-34			85.51	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	1	140	2.8e-33			88.47	ZINC FINGER PROTEIN	COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GLI1; CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	238	377	1.3e-33	0.18	0.95		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	266	433	3.4e-61	0.07	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	2	141	2.8e-33			59.85	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	330	458	2.8e-35	0.16	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	350	489	1.2e-66	0.19	0.96		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	378	544	1e-67	0.13	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	414	545	4.2e-35	0.00	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	442	566	8.4e-33	-0.02	0.87		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

686

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1017	2gli	A	462	563	1.7e-45	0.30	0.78		ZINC FINGER PROTEIN GLI; CHAIN: A; DNA; CHAIN: C, D;	(DNA-BINDING PROTEIN/DNA) COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1017	2gli	A	5	154	7e-32			62.95	ZINC FINGER PROTEIN GLI; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1020	1a7i		137	194	1.7e-16	0.13	0.99		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1020	1a7i		258	313	5.1e-10	0.37	0.66		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1020	1a7i		9	69	5.6e-13			51.43	QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
1020	1b8t	A	138	313	5.1e-29	-0.43	0.15		CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
1020	1b8t	A	199	313	2.8e-11	-0.05	0.28		CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
1020	1b8t	A	74	273	1.7e-37			102.41	CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1020	1b8t	A	76	251	1.7e-37	-0.51	0.57		CRP1; CHAIN: A;	DIFFERENTIATION, CONTRACTILE
1020	1c1l		136	192	1.7e-14	0.19	0.94		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
1020	1c1l		1	79	2.8e-18			50.28	AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
1020	1c1l		258	313	3.4e-10	0.69	0.51		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
1020	1c1l	A	136	194	8.5e-15	0.41	0.81		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1020	1c1l	A	256	313	1.2e-11	0.71	0.99		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1020	1c1l	A	8	67	1.3e-16			53.17	CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1020	1c1l	A	46	210	2.8e-07			62.37	TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
1020	1g44	C	31	290	0.0068			77.40	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1020	1iml		138	203	5.1e-15	0.20	0.18		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
1020	1k1o		14	182	6.8e-07			69.70	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN

688

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1020	1mey	C	221	307	5.6e-36	0.03	-0.15		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	1a1h	A	350	432	3e-28			78.62	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1024	1mey	C	17	93	1.1e-46			61.77	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	1mey	C	181	262	1.1e-47	0.11	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	1mey	C	209	290	4.2e-49	0.23	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	1mey	C	21	103	1.4e-50			99.15	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	1mey	C	237	318	1.1e-49	0.70	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1024	Imey	C	293	374	1.4e-50	0.14	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	Imey	C	29	111	1.4e-50			97.98	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	Imey	C	321	402	1.4e-50	0.32	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	Imey	C	349	430	1.4e-50	0.03	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	Imey	C	349	431	1.4e-50			100.76	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	Imey	C	377	438	5.6e-38	0.06	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	Imey	C		64	7e-40			59.08	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

690

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1024	1tt6	A	1	133	7e-34			64.03	CHAIN: C, F, G; TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1024	1tt6	A	1	144	1.4e-35			95.70	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1024	1tt6	A	265	430	1.5e-70			116.54	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1024	1ubd	C	181	291	3e-52			83.93	YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1024	1ubd	C	189	290	1.4e-35	0.16	-1202.08		YY1; CHAIN: C; ADENOVIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1024	1ubd	C	1	111	2.8e-36			83.50	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1024	1ubd	C	217	318	2.8e-35	0.16	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1024	1ubd	C	21	131	1.4e-36			84.44	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1024	1ubd	C	235	346	3e-52	0.23	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID No.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1024	2gli	A	153	292	2.8e-33	0.01	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	REGULATION/DNA) COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1024	2gli	A	1	132	1.4e-32			87.11	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1024	2gli	A	1	133	8.4e-32			86.13	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1024	2gli	A	209	348	3e-66	0.17	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1024	2gli	A	217	345	4.2e-35	0.24	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1024	2gli	A	237	376	1.3e-67			98.02	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1024	2gli	A	301	429	7e-35	0.01	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1026	1dkc	A	1	222	1.5e-05			51.64	SUBSTRATE BINDING DOMAIN OF DNAK; CHAIN: A; SUBSTRATE PEPTIDE (7 RESIDUES); CHAIN: B;	COMPLEX (MOLECULAR CHAPERONE/PEPTIDE) DNAK, HEAT SHOCK PROTEIN 70 KDA (HSP70), COMPLEX 2 (MOLECULAR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1026	1dky	B	1	206	0.0001			50.66	DNAK; CHAIN: A, B; PEPTIDE SUBSTRATE; CHAIN: C, D;	CHAPERONE/PEPTIDE) COMPLEX (MOLECULAR CHAPERONE/PEPTIDE) DNAK, HEAT SHOCK PROTEIN 70 KDA (HSP70), COMPLEX 2 (MOLECULAR CHAPERONE/PEPTIDE)
1029	1aut	L	261	326	0.00024	0.03	0.33		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
1029	1dva	L	261	326	0.0006	-0.18	0.19		DES-GI A FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GI A FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
1029	1eis	A	115	200	7.2e-09	-0.02	0.11		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
1029	1eis	A	196	270	6e-05	0.15	0.07		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
1029	1ext	A	65	227	3e-27	0.00	-1202.08		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1029	1ext	A	98	249	2.4e-14			63.01	TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
1029	1g40	A	26	268	3e-29			93.82	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
1029	1g44	A	36	268	3e-28			93.40	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1029	1g44	B	42	268	6e-33			90.73	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1029	1g44	C	42	277	9e-24			92.61	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1029	1klo		22	216	3e-14	0.04	-1202.08		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1029	1klo		27	190	3e-14			71.49	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1029	1pfx	L	10	144	2.4e-13			64.18	FACTOR IXA; CHAIN: C; L <sub>2</sub> ; D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR, COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
1029	1qub	A	7	275	3e-30			68.14	HUMAN BETA2-GLYCOPROTEIN I <sub>1</sub> ; CHAIN: A <sub>1</sub>	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
1029	9wga	A	93	269	1.5e-30			94.61	LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1031	lawq	A	2	164	2.8e-91	1.33	-1202.08		(ISOLECTIN 2) 9WGA 3	
									CYCLOPHILIN A; CHAIN: A; PEPTIDE FROM THE HIV-1 CAPSID PROTEIN; CHAIN: B;	COMPLEX (ISOMERASE/PEPTIDE) COMPLEX (ISOMERASE/PEPTIDE), CYCLOPHILIN A, HIV-1 CAPSID, 2 PSEUDO-SYMMETRY
1031	lawq	A	2	165	2.8e-91			269.60	CYCLOPHILIN A; CHAIN: A; PEPTIDE FROM THE HIV-1 CAPSID PROTEIN; CHAIN: B;	COMPLEX (ISOMERASE/PEPTIDE) COMPLEX (ISOMERASE/PEPTIDE), CYCLOPHILIN A, HIV-1 CAPSID, 2 PSEUDO-SYMMETRY
1031	lctq	A	7	178	2.8e-63			98.08	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
1035	lalh	A	13	99	2.8e-29			77.80	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1035	lalh	A	98	182	8.4e-23	0.31	-1202.08		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1035	lmey	C	129	210	2.8e-41	0.38	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	lmey	C	12	94	9.8e-51			95.25	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FINGER PROTEIN; CHAIN: C, F, G;	INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	12	98	4.2e-49			94.62	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	157	238	4.2e-44	0.20	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	185	266	2.8e-46	0.09	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	1	67	1.4e-41			73.50	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	213	294	4.2e-48	0.10	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	23	105	9.8e-51			96.53	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C							DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	Imey	C	297	378	2.8e-49	0.50	-1202.08		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	Imey	C	325	406	7e-50	0.16	-1202.08		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	Imey	C	353	434	7e-50	0.49	-1202.08		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	Imey	C	381	462	2.8e-50	0.27	-1202.08		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	Imey	C	409	490	5.6e-51	0.51	-1202.08		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	Imey	C	437	518	1.4e-50	0.28	-1202.08		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1035	1mev	C	465	546	1.3e-50	0.23	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	493	574	8.4e-50	0.03	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	49	131	1.4e-50			89.69	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	6	85	7e-50			90.19	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	97	182	9.8e-38	0.42	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1mev	C	98	182	3e-38	0.45	-1202.08		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1035	1tt3	A	12	101	4.2e-14			58.28	TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TTIIIA; 5S GENE; NMR, TTIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1035	1tt6	A	1	134	8.4e-36			71.45	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
										COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1035	1tt6	A	1	145	1.4e-35			84.26	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
										COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1035	1tt6	A	242	387	5.6e-38	0.02	-1202.08		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
										COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1035	1tt6	A	381	544	4.5e-69			113.05	TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
										COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1035	1tt6	A	410	556	9.8e-39	0.16	-1202.08		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA
										COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1035	1tt6	A	438	574	1.4e-37	0.07	-1202.08		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1035	1tt6	A	439	574	1e-53	0.16	-1202.08		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1035	1tt6	A	439	574	1e-53	0.16	-1202.08		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1035	1ubd	C	100	210	4.2e-27	0.01	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	12	106	1.4e-29			68.31	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	1	103	4.2e-36			76.28	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG

701

Table 5

SEQ ID No.	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INITIATOR ELEMENT DNA; CHAIN: A, B;	1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	23	133	4.2e-36			83.22	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	323	434	4.5e-49	0.40	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	358	462	1.4e-36	0.04	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	417	518	1.4e-36	0.02	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1035	1ubd	C	435	546	6e-51	0.41	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	463	574	1.5e-49	-0.00	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	1ubd	C	67	182	1.5e-37	0.45	-1202.08		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1035	2gli	A	1	110	1.3e-25			53.64	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	1	132	4.2e-33			76.13	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1035	2gli	A	1	134	2.8e-33			82.37	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	297	436	4.5e-65			98.88	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	297	492	4.5e-65	0.00	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	333	461	7e-35	0.54	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	353	520	7.5e-64	0.04	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	389	520	2.8e-35	0.13	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	417	545	8.4e-35	0.19	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	439	574	1.2e-61	0.16	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1035	2gli	A	69	212	1.1e-26	0.07	-1202.08		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI;

Table 5

SEQ ID No.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: C, D;	GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1042	12e8	H	336	518	5.6e-51	-0.05	0.10		2E8 (IGG1=KAPPA=) ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN IMMUNOGLOBULIN
1042	1a3l	H	336	515	4.2e-50	0.02	0.19		IMMUNOGLOBULIN FAB 13G5; CHAIN: L, H;	IMMUNOGLOBULIN DIEI.S-ALDER, DISFAVORED REACTION, CATALYTIC ANTIBODY, 2 IMMUNOGLOBULIN
1042	1adq	L	51	240	8.4e-32	-0.11	0.18		IGG4 REA; CHAIN: A; RF-ANIGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
1042	1a5v	H	336	517	5.6e-51	0.17	0.10		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN) HIV-1 CA, HIV CA, HIV P24, P24; FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN), HIV, CAPSID PROTEIN, 2 P24
1042	1b1h	A	149	506	2.8e-27	0.13	0.87		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1042	1b1h	A	150	515	8.5e-45	0.37	0.80		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1042	1b1h	A	46	417	3.4e-48			130.68	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1042	1b1h	A	50	417	3.4e-48	0.15	1.00		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION



705

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1042	1bm3	H	336	519	1.3e-50	0.14	0.09		IMMUNOGLOBULIN OPG2 FAB, CONSTANT DOMAIN; CHAIN: L; IMMUNOGLOBULIN OPG2 FAB, VARIABLE DOMAIN; CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN
1042	1c5c	H	336	519	2.8e-51	0.10	0.06		CHIMERIC DECARBOXYLASE ANTIBODY 21D8; CHAIN: L; CHIMERIC DECARBOXYLASE ANTIBODY 21D8; CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN, CATALYTIC ANTIBODY, CHIMERIC FAB, 2 DECARBOXYLASE, HAPTEN COMPLEX
1042	1c1c	B	336	515	5.6e-51	-0.00	-0.02		IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
1042	1c1c	B	48	244	2.8e-72	0.16	-0.06		IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
1042	1cl7	I	434	519	7e-25	0.08	0.09		IGG1 ANTIBODY 1696 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 1696 (VARIABLE HEAVY CHAIN); CHAIN: H; IGG1 ANTIBODY 1696	IMMUNE SYSTEM IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1 FAB FRAGMENT, CROSS-REACTIVITY, HIV1 PROTEASE, ENZYME 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(CONSTANT HEAVY CHAIN); CHAIN: I;	INHIBITION, IMMUNOGLOBULIN
1042	1cql	A	418	516	1.4e-25	0.23	-0.11		CH3 DOMAIN OF MAK33 ANTIBODY; CHAIN: A, B;	IMMUNE SYSTEM CONSTANT DOMAIN, C1-SUBSET, IMMUNOGLOBULIN, IMMUNE SYSTEM
1042	1cs6	A	135	515	3.4e-46	0.31	0.62		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1042	1cs6	A	149	517	1.4e-35	0.19	0.77		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1042	1cs6	A	244	602	1.1e-34	0.11	0.46		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1042	1cs6	A	43	418	1.7e-56			125.44	AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1042	1cs6	A	50	416	1.7e-56	-0.10	0.71		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1042	1dgi	R	151	416	5.1e-28	-0.23	0.47		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1042	1dgi	R	49	331	3.4e-43			112.84	POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1042	1dgi	R	52	331	3.4e-43	-0.28	0.05		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1042	1dh2	A	250	415	1.3e-35	-0.05	0.07		IMMUNOGLOBULIN LAMBDA HEAVY CHAIN; CHAIN: A, B; ENGINEERED PEPTIDE; CHAIN: E, F;	RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1042	1e4k	A	244	415	1.4e-36	-0.26	0.06		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C; FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
1042	1e4x	H	336	518	4.2e-51	0.32	0.27		TAB2; CHAIN: L, M; TAB2; CHAIN: H, I; CYCLIC PEPTIDE; CHAIN: P, Q	COMPLEX (ANTIBODY/ANTIGEN) CROSS-REACTIVITY, PROTEIN-PEPTIDE RECOGNITION
1042	1e4x	H	48	247	1.4e-74	0.08	-0.11		TAB2; CHAIN: L, M; TAB2; CHAIN: H, I; CYCLIC PEPTIDE; CHAIN: P, Q	COMPLEX (ANTIBODY/ANTIGEN) CROSS-REACTIVITY, PROTEIN-PEPTIDE RECOGNITION
1042	1eap	B	49	241	1.1e-65	-0.06	0.29		CATALYTIC ANTIBODY 17B8 COMPLEXED WITH PHENYL [1-(1-N-SUCCINYLAMINO)PENTYL] LEAP 3 PHOSPHONATE LEAP 4	
1042	1eio	H	51	244	1.1e-65	0.05	0.00		IGG2A MONOCLONAL ANTIBODY (LIGHT CHAIN); CHAIN: L; IGG2A MONOCLONAL ANTIBODY (HEAVY CHAIN); CHAIN: H; FMDV PEPTIDE; CHAIN:	IMMUNE SYSTEM FMDV, ANTIGENIC-ANTIBODY INTERACTIONS, RGD MOTIF, G-H LOOP 2 OF VP1.

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1042	1evt	C	247	416	8.5e-27	0.27	0.88		P <sub>1</sub> FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1042	1f2q	A	250	420	1.2e-26	0.13	0.87		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RL-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1042	1f6a	A	246	420	5.1e-31	0.06	0.72		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1042	1fai	H	48	241	8.4e-68	0.07	-0.05		IMMUNOGLOBULIN FAB FRAGMENT FROM A MONOCLONAL ANTILARSONATE ANTIBODY, R19.9 IFAI 3 (IGG2B.KAPPA) IFAI 4	
1042	1fbi	H	48	244	1.1e-72	0.05	-0.17		COMPLEX (ANTIBODY/ANTIGEN) FAB FRAGMENT OF THE MONOCLONAL ANTIBODY P9.13.7 (IGG1) IFBI 3	
									COMPLEXED WITH LYSOZYME (E.C.3.2.1.17) IFBI 4	
1042	1f62	D	250	415	7e-36	0.07	0.06		IMMUNOGLOBULIN IMMUNOGLOBULIN FC	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									AND FRAGMENT B OF PROTEIN A COMPLEX 1FC2 4	
1042	1fgn	H	337	514	7e-51	0.11	0.48		IMMUNOGLOBULIN FAB 5G9; CHAIN: L, H;	IMMUNOGLOBULIN FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN; ANTIBODY, FAB, ANTI-TF, MONOCLONAL, MURINE, IMMUNOGLOBULIN
1042	1fh5	H	54	244	8.4e-66	0.14	0.07		MONOCLONAL ANTIBODY MAK33; CHAIN: L; MONOCLONAL ANTIBODY MAK33; CHAIN: H;	IMMUNE SYSTEM FAB, BIP, CRYSTAL STRUCTURE
1042	1fh3	A	51	243	7e-66	-0.02	0.06		BLUE FLUORESCENT ANTIBODY (19G2)-HEAVY CHAIN; CHAIN: H, A; BLUE FLUORESCENT ANTIBODY (19G2)-LIGHT CHAIN; CHAIN: L, B;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD
1042	1for	H	50	244	4.2e-72	0.03	-0.06		IMMUNOGLOBULIN 1G2A FAB FRAGMENT (FAB17-1A) (ORTHORHOMBIC CRYSTAL FORM) 1FOR 3	
1042	1fsk	C	336	519	1.4e-50	0.11	0.01		MAJOR POLLEN ALLERGEN BET V 1-A; CHAIN: A, D, G, J; IMMUNOGLOBULIN KAPPA LIGHT CHAIN; CHAIN: B, E, H, K; ANTIBODY HEAVY CHAIN FAB; CHAIN: C, F;	IMMUNE SYSTEM BET V 1-A, BETV1 ALLERGEN; BV16 FAB-FRAGMENT, KAPPA MOPC21 CODING SEQUENCE; HEAVY CHAIN OF THE MONOCLONAL ANTIBODY MST2; BET V 1, BV16 FAB FRAGMENT, ANTIBODY ALLERGEN COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQ/OL D score	Compound	PDB annotation
1042	1fsk	C	48	244	2.8e-72	0.05	-0.14		1, I; MAJOR POLLEN ALLERGEN BET V 1-A; CHAIN: A, D, G, J; IMMUNOGLOBULIN KAPPA LIGHT CHAIN; CHAIN: B, E, H, K; ANTIBODY HEAVY CHAIN FAB; CHAIN: C, F, I, L;	IMMUNE SYSTEM BET V 1-A, BETVI ALLERGEN; BV16 FAB-FRAGMENT, KAPPA MOPC21 CODING SEQUENCE; HEAVY CHAIN OF THE MONOCLONAL ANTIBODY MST2, BET V 1, BV16 FAB FRAGMENT, ANTIBODY ALLERGEN COMPLEX
1042	1hi6	B	50	246	1.1e-70	0.29	0.22		IGG2A KAPPA ANTIBODY CB41 (LIGHT CHAIN); CHAIN: A; IGG2A KAPPA ANTIBODY CB41 (HEAVY CHAIN); CHAIN: B; PEPTIDE 5; CHAIN: C; I, L;	COMPLEX (ANTIBODY/PEPTIDE) POLYSPECIFICITY, CROSSREACTIVITY, FAB-FRAGMENT, PEPTIDE, 2 HIV-1
1042	1hzh	H	163	521	4.2e-63	0.12	0.55		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
1042	1hzh	H	48	418	0	0.28	0.55		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
1042	1hzh	H	4	320	1.4e-51	-0.22	0.81		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1042	1i1c	A	251	415	7e-30	0.07	-0.11		IG GAMMA-2A CHAIN C REGION; CHAIN: A, B;	IMMUNE SYSTEM IGG2A; IGG, FC
1042	1i1g	H	49	241	5.6e-67	0.10	0.13		IMMUNOGLOBULIN IGG FAB (GGG2B, KAPPA) FRAGMENT (40-50 FAB) COMPLEXED WITH IIBG 3 OULABAIN IIBG 4	
1042	1i1g	B	160	517	9.8e-61	-0.03	0.54		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1042	1i1g	B	49	415	0	0.10	0.66		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1042	1i1g	B	51	475	0			102.57	IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1042	1i1g	B	5	318	2.8e-48	-0.12	0.28		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1042	1i1g	B	160	516	5.6e-61	0.04	0.64		IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION, C REGION, HINGE REGION
1042	1i1g	B	49	415	0	-0.03	0.31		IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
1042	1i1i	A	50	244	7e-68	0.06	-0.02		MONOCLONAL ANTIBODY G3-519 (HEAVY CHAIN); CHAIN: A; MONOCLONAL ANTIBODY G3-519 (LIGHT CHAIN); CHAIN: B;	IMMUNE SYSTEM FAB, BETA SHEET STRUCTURE, ANTIBODY
1042	1i1b	B	261	517	1e-36	0.15	0.71		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1	COMPLEX (IMMUNOGLOBULIN/RECEPTOR)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INTERLEUKIN-1 RECEPTOR; CHAIN: B;	IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1042	1kb5	H	49	244	2.8e-72	0.10	0.06		KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
1042	Imco	H	189	516	1.1e-67	0.05	0.01		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1042	Imco	H	48	415	0	0.05	0.66		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1042	Imco	H	49	475	0			106.26	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1042	Imco	H	5	320	2.8e-55	-0.30	0.09		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1042	Imcp	H	49	231	1.1e-48	-0.07	0.11		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB FRAGMENT (MCP/C\$603)	



713

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1042	1mlb	B	336	519	1.4e-50	0.16	0.07		IMCP 4 IMMUNOGLOBULIN FAB D44.1 (GG1,KAPPA) (BALB/C MOUSE, MONOCLONAL ANTIBODY) 1MLB 5	
1042	1ptc		417	520	2.8e-23	0.02	-0.13		IMMUNOGLOBULIN \$P/\$\$C(PRIME) FRAGMENT OF AN IG*G1 1PFC 4	
1042	1plg	H	48	243	4.2e-73	0.17	-0.05		IGG2A=KAPPA=; 1PLG 4 CHAIN: L, H; 1PLG 5	IMMUNOGLOBULIN
1042	1psk	H	49	240	1.4e-61	-0.17	0.04		ANTIBODY; CHAIN: L, H;	IMMUNOGLOBULIN FAB, GD2-GANGLIOSIDE, CARBOHYDRATE, MELANOMA, IMMUNOGLOBULIN
1042	1qfu	H	336	518	1.3e-50	0.30	0.28		HEMAGGLUTININ (HA1 CHAIN); CHAIN: A;	VIRAL PROTEIN/IMMUNE SYSTEM COMPLEX
									HEMAGGLUTININ (HA2 CHAIN); CHAIN: B;	(HEMAGGLUTININ/IMMUNOGLOBULIN), HEMAGGLUTININ, 2
									IMMUNOGLOBULIN	IMMUNOGLOBULIN, VIRAL
									IGG1-KAPPA ANTIBODY (LIGHT CHAIN); CHAIN: L; IMMUNOGLOBULIN	PROTEIN/IMMUNE SYSTEM
									IGG1-KAPPA ANTIBODY (HEAVY CHAIN); CHAIN: H;	
1042	1qfu	H	48	244	1.4e-72	0.10	-0.11		HEMAGGLUTININ (HA1 CHAIN); CHAIN: A;	VIRAL PROTEIN/IMMUNE SYSTEM COMPLEX
									HEMAGGLUTININ (HA2 CHAIN); CHAIN: B;	(HEMAGGLUTININ/IMMUNOGLOBULIN), HEMAGGLUTININ, 2
									IMMUNOGLOBULIN	IMMUNOGLOBULIN, VIRAL
									IGG1-KAPPA ANTIBODY (LIGHT CHAIN); CHAIN: L; IMMUNOGLOBULIN	PROTEIN/IMMUNE SYSTEM
									IGG1-KAPPA ANTIBODY	

714

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(HEAVY CHAIN); CHAIN: H <sub>1</sub>	
1042	1vge	H	51	244	1.4e-66	-0.10	0.18		TR1.9 FAB; CHAIN: L <sub>1</sub> , H <sub>1</sub>	IMMUNOGLOBULIN TR1.9, ANTI-THYROID PEROXIDASE, AUTOANTIBODY, 2
									IMMUNOGLOBULIN	
1042	1wej	H	336	520	8.4e-51	0.24	0.27		E8 ANTIBODY; CHAIN: L <sub>1</sub> , H <sub>1</sub> , CYTOCHROME C <sub>1</sub> ; CHAIN: F <sub>1</sub>	COMPLEX (ANTIBODY/ELECTRON TRANSPORT) FAB E8; CYT C <sub>1</sub> , ANTIGEN; IMMUNOGLOBULIN, IGG1 KAPPA, FAB FRAGMENT, HORSE 2 CYTOCHROME C <sub>1</sub> , COMPLEX (ANTIBODY/ELECTRON TRANSPORT)
									IG ANTIBODY D2.3 (LIGHT CHAIN); CHAIN: L <sub>1</sub> ; IG ANTIBODY D2.3 (HEAVY CHAIN); CHAIN: H <sub>1</sub>	IMMUNE SYSTEM ABZYME, TRANSITION STATE ANALOG, IMMUNE SYSTEM
1042	25c8	H	336	515	1.4e-51	0.11	0.16		IGG 5C8; CHAIN: L <sub>1</sub> , H <sub>1</sub>	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
									IMMUNOGLOBULIN IG* A FAB FRAGMENT (J539) (GALACTAN-BINDING) 2FBJ 3	
1042	2fbj	H	49	231	7e-51	-0.03	0.04		FC GAMMA RIIB; CHAIN: A <sub>1</sub>	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1042	2fcb	A	249	419	3.4e-28	0.11	0.74		METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: A <sub>1</sub> , C <sub>1</sub> , METAL CHELATASE	IMMUNE SYSTEM METAL CHELATASE, CATALYTIC ANTIBODY, FAB FRAGMENT, IMMUNE 2 SYSTEM
1042	3fct	B	336	518	4.2e-50	0.22	0.09			

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1043	1cdq		88	164	8.4e-20	0.59	1.00		CATALYTIC ANTIBODY; CHAIN: B, D; COMPLEMENT REGULATORY PROTEIN CD59 (NMR, 20 STRUCTURES) 1CDQ 3	
1043	1cdq		88	164	8.4e-20			142.11	COMPLEMENT REGULATORY PROTEIN CD59 (NMR, 20 STRUCTURES) 1CDQ 3	
1043	1erg		88	157	2.8e-19	0.41	1.00		COMPLEMENT FACTOR HUMAN COMPLEMENT REGULATORY PROTEIN CD59 (EXTRACELLULAR 1ERG 3 REGION, RESIDUES 1 - 70) (NMR, RESTRAINED MINIMIZED 1ERG 4 AVERAGE STRUCTURE) 1ERG 5	
1043	1erg		88	157	2.8e-19			131.94	COMPLEMENT FACTOR HUMAN COMPLEMENT REGULATORY PROTEIN CD59 (EXTRACELLULAR 1ERG 3 REGION, RESIDUES 1 - 70) (NMR, RESTRAINED MINIMIZED 1ERG 4 AVERAGE STRUCTURE) 1ERG 5	
1047	1aln	A	11	199	0			232.06	B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C; COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1047	1aln	A	25	299	0			410.38	B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	(ANTIGEN/PEPTIDE) COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1047	1aln	A	25	300	0	0.83	1.00		B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1047	1agd	A	11	199	0			233.86	B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL - INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1047	1agd	A	25	299	0			411.64	B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL - INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1047	1agd	A	25	300	0	0.79	1.00		B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL - INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1047	1efx	A	25	302	0	0.87	1.00		HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX

717

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	
1047	1efx	A	25	302	0			420.37	HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1047	1hsa	A	11	199	0			232.10	HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN I HSA 3 /HLA-B(ASTERISK)2705\$ I HSA 4	
1047	1hsa	A	25	299	0			410.78	HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN I HSA 3 /HLA-B(ASTERISK)2705\$ I HSA 4	
1047	1hsa	A	25	300	0	0.82	1.00		HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN I HSA 3 /HLA-	

718

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									B(ASTERISK)2705\$ 1HSA 4	
1047	1hsb	A	11	199	0			257.74	HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE 1HSB 3 ANTIGEN) 1HSB 4	
1047	1hsb	A	25	294	0	0.86	1.00		HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE 1HSB 3 ANTIGEN) 1HSB 4	
1047	1hsb	A	25	294	0			414.03	HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE 1HSB 3 ANTIGEN) 1HSB 4	
1047	1i4f	A	11	199	1.4e-100			263.59	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B: MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUCOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1047	1i4f	A	25	299	0	0.83	1.00		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B: MELANOMA-ASSOCIATED ANTIGEN	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUCOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1047	1i4f	A	25	299	0			435.50	4; CHAIN: C <sub>2</sub> HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A <sub>2</sub> ; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C <sub>2</sub>	IMMUNE SYSTEM IMAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1047	1qgd	A	12	199	0			232.33	HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A <sub>2</sub> ; BETA-2-MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C <sub>2</sub>	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1047	1qgd	A	26	298	0	0.87	1.00		HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A <sub>2</sub> ; BETA-2-MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C <sub>2</sub>	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1047	1qgd	A	26	298	0			407.96	HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A <sub>2</sub> ; BETA-2-MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C <sub>2</sub>	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1047	1tmc	A	11	185	2.8e-94			284.63	HISTOCOMPATIBILITY ANTIGEN TRUNCATED	

720

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1048	1efx	A	11	199	0			241.21	HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-A*W68 ITMC 3 COMPLEXED WITH A DECAMERIC PEPTIDE (EVAPPEYHRK) ITMC 4	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1048	2ebo	A	12	88	4.2e-18	-0.74	0.29		EBOLA VIRUS ENVELOPE GLYCOPROTEIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D; E;	ENVELOPE GLYCOPROTEIN ENVELOPE GLYCOPROTEIN, FILOVIRUS, EBOLA VIRUS, GP2, COAT 2 PROTEIN
1050	1cdy		32	133	3.4e-07	0.40	0.19		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
1050	1dgi	R	25	127	2.8e-29	0.41	0.46		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR



721

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1050	1dt9	A	32	126	8.5e-10	0.33	0.06		T LYMPHOCYTE ACTIVATION ANTIGEN; CHAIN: A;	IMMUNE SYSTEM B7-1 (CD80); IG SUPERFAMILY
1050	1eaj	A	33	127	1.7e-07	0.41	-0.03		COXSACKIE VIRUS AND ADENOVIRUS RECEPTOR; CHAIN: A, B;	VIRUS/VIRAL PROTEIN RECEPTOR COXSACKIEVIRUS B-ADENOVIRUS RECEPTOR, HCAR, VIRUS/VIRAL PROTEIN RECEPTOR, IMMUNOGLOBULIN V DOMAIN FOLD, 2 SYMMETRIC DIMER
1050	1hxm	B	20	123	1.5e-06	0.36	0.06		GAMMA-DELTA T-CELL RECEPTOR; CHAIN: A, C, E, G; GAMMA-DELTA T-CELL RECEPTOR; CHAIN: B, D, F, H;	IMMUNE SYSTEM T-CELL RECEPTOR DELTA CHAIN; T-CELL RECEPTOR GAMMA CHAIN; IG DOMAIN, T CELL RECEPTOR, TCR, GPCR
1050	1i8l	C	33	120	8.5e-06	0.12	0.80		T LYMPHOCYTE ACTIVATION ANTIGEN CD80; CHAIN: A, B; CYTOTOXIC T-LYMPHOCYTE PROTEIN 4; CHAIN: C, D;	IMMUNE SYSTEM ACTIVATION B7-1 ANTIGEN, CTLA-4 COUNTER-RECEPTOR CTLA-4, CYTOTOXIC T-LYMPHOCYTE-ASSOCIATED ANTIGEN RECEPTORS, INHIBITORY COMPLEX
1050	1iil	G	32	131	7e-06	0.21	-0.03		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
1050	1neu		27	128	6.8e-11	0.56	0.13		MYELIN P0 PROTEIN; CHAIN: NULL;	STRUCTURAL PROTEIN MYELIN, STRUCTURAL PROTEIN, GLYCOPROTEIN, TRANSMEMBRANE, PHOSPHORYLATION, IMMUNOGLOBULIN FOLD, SIGNAL, MYELIN 2 MEMBRANE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1050	2cd0	A	29	127	8.4e-06	0.35	0.31		BENCE-JONES PROTEIN WIL, A VARIABLE DOMAIN FROM CHAIN: A, B;	ADHESION MOLECULE
1052	1tgy	A	55	98	0.0031	-0.49	0.01		CYTOTOXIN TOXIN GAMMA (CARDIOTOXIN) ITGX 3	
1052	2crs		55	98	0.0023	-0.25	0.00		CARDIOTOXIN CARDIOTOXIN III (NMR, 13 STRUCTURES) 2CRS 3	
1054	1a5e		117	272	8.4e-20			74.90	TUMOR SUPPRESSOR P16INK4A; CHAIN: NULL;	ANTI-ONCOGENE CELL CYCLE, ANTI-ONCOGENE, REPEAT, ANK REPEAT
1054	1awc	B	95	244	7e-38			75.28	GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
1054	1bd8		93	247	4.2e-31			71.99	P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
1054	1bx	B	95	250	2.8e-31			71.23	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
1054	1bu9	A	91	255	4.2e-33			80.55	CYCLIN-DEPENDENT	HORMONE/GROWTH FACTOR

723

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									KINASE 6 INHIBITOR; CHAIN: A;	P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
1054	1by2		1	113	8.4e-44			114.47	MAC-2 BINDING PROTEIN; CHAIN: NULL;	EXTRACELLULAR MODULE TUMOR-ASSOCIATED ANTIGEN 90K; EXTRACELLULAR MODULE, SCAVENGER RECEPTOR, TUMOUR-ASSOCIATED 2 ANTIGEN, EXTRACELLULAR MATRIX, GLYCOSYLATED PROTEIN
1054	1by2		711	824	7e-44			113.40	MAC-2 BINDING PROTEIN; CHAIN: NULL;	EXTRACELLULAR MODULE TUMOR-ASSOCIATED ANTIGEN 90K; EXTRACELLULAR MODULE, SCAVENGER RECEPTOR, TUMOUR-ASSOCIATED 2 ANTIGEN, EXTRACELLULAR MATRIX, GLYCOSYLATED PROTEIN
1054	1by2		714	822	7e-44	0.79	1.00		MAC-2 BINDING PROTEIN; CHAIN: NULL;	EXTRACELLULAR MODULE TUMOR-ASSOCIATED ANTIGEN 90K; EXTRACELLULAR MODULE, SCAVENGER RECEPTOR, TUMOUR-ASSOCIATED 2 ANTIGEN, EXTRACELLULAR MATRIX, GLYCOSYLATED PROTEIN
1054	1ctu	A	217	709	1.4e-74			146.51	SOLUBLE QUINOPROTEIN GLUCOSE DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE BETA-PROPELLER, SUPERBARREL, COMPLEX WITH THE COFACTOR PQQ 2 AND THE INHIBITOR METHYLHYDRAZINE, OXIDOREDUCTASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1054	1ctu	A	218	645	1.4e-74	0.34	0.92		SOLUBLE QUINOPROTEIN GLUCOSE DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE BETA-PROPELLER, SUPERBARREL, COMPLEX WITH THE COFACTOR PQQ 2 AND THE INHIBITOR METHYLHYDRAZINE, OXIDOREDUCTASE
1054	1d9s	A	2	129	2.8e-07			51.72	CYCLIN-DEPENDENT KINASE 4 INHIBITOR B; CHAIN: A;	SIGNALING PROTEIN HELIX-TURN-HELIX, ANKYRIN REPEAT
1054	1ihb	A	96	246	4.2e-33			78.44	CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
1054	1ikm	D	95	296	2.8e-38			80.27	NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFkB COMPLEX
1054	1myo		127	244	1.3e-26			72.80	MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
1054	1nfi	E	87	292	5.6e-38			75.42	NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
1057	1e9t	A	220	269	1.3e-11	0.15	-1202.08		INTESTINAL TREFOIL FACTOR; CHAIN: A;	CELL MOTILITY FACTOR HITF; INTESTINAL TREFOIL FACTOR, SOLUTION STRUCTURE, TREFOIL 2 DOMAIN, NMR SPECTROSCOPY, CELL MOTILITY FACTOR
1057	1hi7	A	222	275	7.5e-16	0.26	-1202.08		PS2 PROTEIN; CHAIN: A;	GROWTH FACTOR PNR-

725

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									B <sub>1</sub>	2.PS2.TFF1.BREAST CANCER ESTROGEN INDUCIBLE GROWTH FACTOR, CELL MOTILITY, TUMOR SUPPRESSOR, TREFOIL 2 DOMAIN, SIGNAL
1057	2psp	A	223	269	1.5e-11	0.35	-1202.08		PORCINE PANCREATIC SPASMOLYTIC POLYPEPTIDE; CHAIN: A, B;	TREFOIL FAMILY OF PEPTIDES PSP REPEAT, GROWTH FACTOR, SIGNAL
1063	1aln	A	29	255	0	0.27	-1202.08		B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1063	1a6z	A	22	227	1.1e-68			53.05	HFE; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D	MHC CLASS I COMPLEX HFE, HEREDITARY HEMOCHROMATOSIS, MHC CLASS I
1063	1a6z	A	29	255	1.4e-68			59.25	HFE; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D	MHC CLASS I COMPLEX HFE, HEREDITARY HEMOCHROMATOSIS, MHC CLASS I
1063	1agd	A	29	255	0	0.36	-1202.08		B*0801; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKKYL-INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1063	1e16	A	29	255	2.8e-67			50.68	MHC-LIKE PROTEIN T22; CHAIN: A, C, E, G; BETA-2-MICROGLOBULIN; CHAIN: B, D, F, H	IMMUNE SYSTEM/NON-CLASSICAL MHC-LIKE, MAJOR HISTOCOMPATIBILITY, BETA2-2 MICROGLOBULIN
1063	1d2v	C	118	585	0			522.22	MYELOPEROXIDASE;	OXIDOREDUCTASE HEME-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQPOL ID score	Compound	PDB annotation
1063	1ed3	A	29	255	0	0.40	-1202.08		CHAIN: A, B; MYELOPEROXIDASE; CHAIN: C, D;	PROTEIN, PEROXIDASE, OXIDOREDUCTASE, PEROXIDASE-2 BROMIDE COMPLEX
										IMMUNE SYSTEM MAJOR HISTOCOMPATIBILITY COMPLEX, RAT MINOR 2 HISTOCOMPATIBILITY COMPLEX, MHC, IMMUNOLOGY, PEPTIDE 3 ANTIGEN PRESENTATION, CELLULAR IMMUNITY, CELL SURFACE 4 RECEPTOR, T CELL RECEPTOR LIGAND
1063	1ed3	A	29	255	0			58.49	CLASS I MAJOR HISTOCOMPATIBILITY ANTIGEN RT1-AA; CHAIN: A, D; BETA-2- MICROGLOBULIN; CHAIN: B, E; PEPTIDE MTF-E (13N3E); CHAIN: C, F;	IMMUNE SYSTEM MAJOR HISTOCOMPATIBILITY COMPLEX, RAT MINOR 2 HISTOCOMPATIBILITY COMPLEX, MHC, IMMUNOLOGY, PEPTIDE 3 ANTIGEN PRESENTATION, CELLULAR IMMUNITY, CELL SURFACE 4 RECEPTOR, T CELL RECEPTOR LIGAND
1063	1efx	A	29	255	0	0.44	-1202.08		HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2- MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1063	1efx	A	29	255	0			57.25	HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	FOLD, RECEPTOR/MHC COMPLEX
1063	1fzk	A	29	255	5.6e-98			58.77	H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, K-B CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; NUCLEOCAPSID PROTEIN; CHAIN: P;	IMMUNE SYSTEM SEV9, MAJOR HISTOCOMPATIBILITY COMPLEX PEPTIDE-MHC
1063	1hoc	A	29	255	2.8e-98			50.16	HISTOCOMPATIBILITY ANTIGEN MURINE CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX CONSISTING 1HOC 3 OF H-2D=B=, B2-MICROGLOBULIN, AND A 9-RESIDUE PEPTIDE 1HOC 4	
1063	1hsa	A	29	255	0	0.24	-1202.08		HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN 1HSA 3 /HLA-B(ASTERISK)2705\$ 1HSA 4	
1063	1hsh	A	29	255	0	0.40	-1202.08		HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1063	1hsb	A	29	255	0			52.66	(LEUCOCYTE IHSB 3 ANTIGEN) IHSB 4	
									HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE IHSB 3 ANTIGEN) IHSB 4	
1063	1hyr	C	21	227	8.4e-55			52.26	NRG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-RELATED PROTEIN A; CHAIN: C;	IMMUNE SYSTEM NKG2D: MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-I, COMPLEX, IMMUNE SYSTEM
1063	1hyr	C	28	255	1.4e-53			67.13	NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-RELATED PROTEIN A; CHAIN: C;	IMMUNE SYSTEM NKG2D: MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-I, COMPLEX, IMMUNE SYSTEM
1063	1i4f	A	29	255	0	0.47	-1202.08		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1063	1i4f	A	29	255	0			60.36	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1063	1ld9	A	29	254	0			56.90	4; CHAIN: C; MHC CLASS I H-2LD HEAVY CHAIN; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; NANO- PEPTIDE; CHAIN: C;	MAJOR HISTOCOMPATIBILITY COMPLEX LD; MAJOR HISTOCOMPATIBILITY COMPLEX, LD
1063	1ld9	A	29	255	0	0.20	-1202.08		MHC CLASS I H-2LD HEAVY CHAIN; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; NANO- PEPTIDE; CHAIN: C;	MAJOR HISTOCOMPATIBILITY COMPLEX LD; MAJOR HISTOCOMPATIBILITY COMPLEX, LD
1063	1qo3	A	30	255	0	0.46	-1202.08		MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
1063	1qo3	A	30	255	0			54.24	MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
1063	1qpd	A	30	255	0	0.20	-1202.08		HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	
1063	1qgd	A	30	255	0			53.86	HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1063	1tmc	A	22	192	9.8e-79			68.60	HISTOCOMPATIBILITY ANTIGEN TRUNCATED HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-A*W68 ITMC 3 COMPLEXED WITH A DECAMERIC PEPTIDE (EVAPPEYHRK) ITMC 4	
1063	1zag	A	29	255	5.6e-62			55.36	ZINC-ALPHA-2-GLYCOPROTEIN; CHAIN: A, B, C, D;	LIPID MOBILIZATION FACTOR ZN-ALPHA-2-GLYCOPROTEIN, ZAG LIPID MOBILIZATION FACTOR, SECRETED MHC CLASS I HOMOLOG
1065	1eqj	A	71	336	4.2e-29	0.11	-1202.08		PHOSPHOGLYCERATE MUTASE; CHAIN: A;	ISOMERASE ALPHA/BETA-TYPE STRUCTURE
1066	12e8	H	2	227	8.4e-09			59.62	2E8 (GG1=KAPPA=) ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN IMMUNOGLOBULIN
1066	1bth	A	29	376	1.1e-29			78.69	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION

731

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1066	1b1h	A	30	346	1.1e-29	0.07	-1202.08		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1066	1cs6	A	20	376	2.8e-35			85.17	AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1066	1cvs	D	122	302	7e-31	0.12	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1066	1cvs	D	37	208	2.8e-23	0.25	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1066	1dgi	R	12	303	9e-22			63.55	POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1066	1ev2	G	132	308	4.2e-30	0.09	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2, FGFR2, IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1066	1eyt	C	37	208	2.8e-22	0.16	-1202.08		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1, FGFR1; IMMUNOGLOBULIN (IG) LIKE

732

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR RECEPTOR 1; CHAIN: C, D;	DOMAINS BELONGING TO THE 1-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1066	1f2q	A	121	308	8.4e-23	0.18	-1202.08		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RL-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1066	1f2q	A	26	214	4.2e-28			66.45	HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RL-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1066	1f2q	A	28	214	4.2e-28	0.23	-1202.08		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RL-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1066	1f42	A	21	325	1.2e-08			63.76	INTERLEUKIN-12 BETA CHAIN; CHAIN: A;	CYTOKINE CYTOKINE
1066	1f6a	A	24	213	1.4e-29			75.69	HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IGE EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1066	1f6a	A	24	214	1.4e-29	0.42	-1202.08		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IGE EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1066	1f8t	H	2	227	4.2e-08			66.69	ANTIBODY FAB FRAGMENT (LIGHT	IMMUNE SYSTEM MONOCLONAL ANTIBODY, ANTIGEN-BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: L; ANTIBODY FAB FRAGMENT (HEAVY CHAIN); CHAIN: H	FRAGMENT, INTERLEUKIN-22, X-RAY ANALYSIS, CRYSTAL
1066	1f97	A	101	309	1.4e-29			72.64	JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1066	1fcg	A	120	305	3e-23	0.14	-1202.08		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1066	1fcg	A	23	210	8.4e-28			78.62	FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1066	1fhl	A	117	307	1.5e-23	0.05	-1202.08		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1066	1fhl	A	22	211	7e-27			73.09	LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1066	1fhl	A	28	212	7e-27	0.17	-1202.08		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1066	1g0x	A	118	310	2.8e-22			72.43	LEUCOCYTE IMMUNOGLOBULIN- LIKE RECEPTOR-1; CHAIN: A;	IMMUNE SYSTEM LEUKOCYTE INHIBITORY RECEPTOR-1; LEUKOCYTE IMMUNOGLOBULIN FOLD, 3-10 HELIX
1066	1g0x	A	120	297	9e-22	0.18	-1202.08		LEUCOCYTE IMMUNOGLOBULIN- INHIBITORY RECEPTOR-1;	IMMUNE SYSTEM LEUKOCYTE INHIBITORY RECEPTOR-1;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1066	1g0x	A	120	306	2.8e-22	0.21	-1202.08		LIKE RECEPTOR-1; CHAIN: A;	LEUKOCYTE IMMUNOGLOBULIN FOLD, 3-10 HELIX
									LEUCOCYTE IMMUNOGLOBULIN-LIKE RECEPTOR-1; CHAIN: A;	IMMUNE SYSTEM LEUKOCYTE INHIBITOR V RECEPTOR-1; LEUKOCYTE IMMUNOGLOBULIN FOLD, 3-10 HELIX
1066	1g0x	A	28	210	5.6e-26	0.21	-1202.08		LEUCOCYTE IMMUNOGLOBULIN-LIKE RECEPTOR-1; CHAIN: A;	IMMUNE SYSTEM LEUKOCYTE INHIBITORY RECEPTOR-1; LEUKOCYTE IMMUNOGLOBULIN FOLD, 3-10 HELIX
1066	1igy	B	3	376	8.4e-09			65.38	IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
1066	Imco	H	2	376	5.6e-10			74.55	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1066	Imkr		29	211	9.8e-26	0.21	-1202.08		P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
1066	Imkr		31	211	5.6e-33			71.40	P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
1066	2dli	A	119	304	2.8e-32	0.16	-1202.08		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2

735

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1066	2dli	A	216	331	9.8e-09	0.06	-1202.08		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNOGLOBULIN
1066	2dli	A	29	210	1.1e-24	0.43	-1202.08		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
1066	2dli	A	31	213	2.8e-32			74.62	MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
1066	2fcb	A	120	306	3e-23	0.14	-1202.08		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1066	2fcb	A	23	214	1.4e-29			81.15	FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1066	2fcb	A	24	213	1.4e-29	0.12	-1202.08		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1066	2nmb	A	1	141	5.6e-33			52.09	NLMB PROTEIN; CHAIN: A; GPVY PEPTIDE; CHAIN: B;	CELL CYCLE/GENE REGULATION COMPLEX, SIGNAL TRANSDUCTION, PHOSPHOTYROSINE BINDING 2 DOMAIN (PTB), ASYMETRIC CELL DIVISION
1066	32c2	B	2	225	9.8e-09			60.00	IGG1 ANTIBODY 32C2; CHAIN: A; IGG1 ANTIBODY 32C2;	IMMUNE SYSTEM FAB, ANTIBODY, AROMATASE, P450

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1067	1hx2	A	8	64	2.8e-16			54.02	CHAIN: B; BSTT; CHAIN: A;	HYDROLASE INHIBITOR BOMBINA SKIN TRYPSIN INHIBITOR BETA-SHEET DISULFIDE-RICH
1069	1b6c	B	1	299	0			340.18	FK506-BINDING PROTEIN; CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
1069	1b6c	B	253	586	0			351.53	FK506-BINDING PROTEIN; CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
1069	1b6c	B	264	581	0	0.64	-1202.08		FK506-BINDING PROTEIN; CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
1069	1es7	B	33	108	1.4e-12	0.14	-1202.08		BONE MORPHOGENETIC PROTEIN-2; CHAIN: A, C; BONE MORPHOGENETIC PROTEIN RECEPTOR 1A; CHAIN: B, D;	CYTOKINE BMP-2; ALK-3; PROTEIN-PROTEIN COMPLEX, THREE FINGER TOXIN FOLD, RECEPTOR-2 LIGAND COMPLEX, CYTOKINE RECEPTOR, TGF BETA SUPERFAMILY
1071	1aln	A	24	299	0			166.47	B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN,



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1071	1aln	A	26	298	0	0.49	-1202.08		B*3501; CHAIN: A, B; PEPTIDE VPLRPMITY; CHAIN: C;	MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1071	1agd	A	24	299	0			169.42	B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL-INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1071	1agd	A	26	298	0	0.41	-1202.08		B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL-INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1071	1c16	A	24	299	5.6e-89			139.27	MHC-LIKE PROTEIN T22; CHAIN: A, C, E, G; BETA-2-MICROGLOBULIN; CHAIN: B, D, F, H	IMMUNE SYSTEM NON-CLASSICAL MHC-LIKE, MAJOR HISTOCOMPATIBILITY, BETA2-2 MICROGLOBULIN
1071	1ed3	A	24	300	0			157.88	CLASS I MAJOR HISTOCOMPATIBILITY ANTIGEN RT1-AA; CHAIN: A, D; BETA-2-MICROGLOBULIN; CHAIN: B, E; PEPTIDE MTF-E (13N3E); CHAIN: C, F;	IMMUNE SYSTEM MAJOR HISTOCOMPATIBILITY COMPLEX, RAT MINOR 2 HISTOCOMPATIBILITY COMPLEX, MHC, IMMUNOLOGY, PEPTIDE 3 ANTIGEN PRESENTATION, CELLULAR IMMUNITY, CELL SURFACE 4 RECEPTOR, T CELL RECEPTOR LIGAND
1071	1ed3	A	26	298	0	0.52	-1202.08		CLASS I MAJOR	IMMUNE SYSTEM MAJOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									HISTOCOMPATIBILITY ANTIGEN RT1-AA; CHAIN: A, D; BETA-2-MICROGLOBULIN; CHAIN: B, E; PEPTIDE MTF-E (13N3E); CHAIN: C, F;	HISTOCOMPATIBILITY COMPLEX, RA1 MINOR 2 HISTOCOMPATIBILITY COMPLEX, MHC, IMMUNOLOGY, PEPTIDE 3 ANTIGEN PRESENTATION, CELLULAR IMMUNITY, CELL SURFACE 4 RECEPTOR, T CELL RECEPTOR LIGAND
1071	1efx	A	24	300	0			163.46	HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1071	1efx	A	26	298	0	0.64	-1202.08		HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1071	1fzk	A	24	296	0			167.88	H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, K-B CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; NUCLEOCAPSID	IMMUNE SYSTEM SEV9; MAJOR HISTOCOMPATIBILITY COMPLEX PEPTIDE-MHC

Table 5

SEQ ID NO.	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1071	1hoc	A	24	295	0			172.82	PROTEIN: CHAIN: P; HISTOCOMPATIBILITY ANTIGEN MURINE CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX CONSISTING 1HOC 3 OF H-2D=B=, B2-MICROGLOBULIN, AND A 9-RESIDUE PEPTIDE 1HOC 4	
1071	1hsa	A	24	299	0			167.79	HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN 1HSA 3 /HLA- B(ASTERISK)2705\$ 1HSA 4	
1071	1hsa	A	26	298	0	0.49	-1202.08		HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN 1HSA 3 /HLA- B(ASTERISK)2705\$ 1HSA 4	
1071	1hsb	A	24	293	0			166.59	HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE 1HSB 3 ANTIGEN) 1HSB 4	
1071	1hsb	A	26	293	0	0.46	-1202.08		HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE 1HSB 3 ANTIGEN) 1HSB 4	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1071	1hyr	C	10	193	2.8e-53			236.94	ANTIGEN I HSB 4	IMMUNE SYSTEM NKG2D, MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-I, COMPLEX, IMMUNE SYSTEM
1071	1hyr	C	23	297	6e-93			439.85	NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-RELATED PROTEIN A; CHAIN: C;	IMMUNE SYSTEM NKG2D; MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-I, COMPLEX, IMMUNE SYSTEM
1071	1i4f	A	24	298	0			168.36	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1071	1i4f	A	26	298	0	0.49	-1202.08		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1071	1i49	A	24	291	0			166.41	MHC CLASS I H-2LD HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; NANO-PEPTIDE; CHAIN: C;	MAJOR HISTOCOMPATIBILITY COMPLEX LD; MAJOR HISTOCOMPATIBILITY COMPLEX, LD

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1071	1mhc	A	24	299	0			156.53	MHC CLASS I ANTIGEN H2-M3; 1MHC 6 CHAIN: A, B, D, E; 1MHC 7 NONAPEPTIDE FROM RAT NADH DEHYDROGENASE; 1MHC 12 CHAIN: C, F; 1MHC 13	HISTOCOMPATIBILITY ANTIGEN/PEPTIDE MAJOR HISTOCOMPATIBILITY COMPLEX; 1MHC 8 ND1, 1MHC 15
1071	1mhc	A	25	296	0			170.11	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-E; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D; PEPTIDE (YMAPRTVL); CHAIN: P, Q;	MAJOR HISTOCOMPATIBILITY COMPLEX MHC NONCLASSICAL CHAIN, MHC-E, HLA-E, MHC CLASS HLA-E, HLA E, MAJOR HISTOCOMPATIBILITY COMPLEX, MHC, HLA, 2 BETA 2 MICROGLOBULIN, PEPTIDE, LEADER PEPTIDE, 3 NON-CLASSICAL MHC, CLASS IB MHC
1071	1mhc	A	26	297	0	0.53	-1202.08		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-E; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D; PEPTIDE (YMAPRTVL); CHAIN: P, Q;	MAJOR HISTOCOMPATIBILITY COMPLEX MHC NONCLASSICAL CHAIN, MHC-E, HLA-E, MHC CLASS HLA-E, HLA E, MAJOR HISTOCOMPATIBILITY COMPLEX, MHC, HLA, 2 BETA 2 MICROGLOBULIN, PEPTIDE, LEADER PEPTIDE, 3 NON-CLASSICAL MHC, CLASS IB MHC
1071	1qo3	A	25	298	0			177.20	MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49

742

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1071	1qo3	A	26	298	0	0.42	-1202.08		MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
1071	1qpd	A	25	296	0			173.19	HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1071	1qpd	A	26	297	0	0.43	-1202.08		HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1071	1tmc	A	11	185	8.4e-80			83.62	HISTOCOMPATIBILITY ANTIGEN TRUNCATED HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-A*W68 1TMC 3 COMPLEXED WITH A DECAMERIC PEPTIDE (EVAPPEYHRK) 1TMC 4	
1071	2fb4	H	212	305	1.1e-07	0.84	-1202.08		IMMUNOGLOBULIN	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQ ID score	Compound	PDB annotation
									IMMUNOGLOBULIN FAB 2FB4 4	
1071	2fgw	H	186	305	4.2e-08	0.18	-1202.08		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
1087	1aox	A	356	548	4.2e-32	0.32	0.96		INTEGRIN ALPHA 2 BETA; CHAIN: A, B;	INTEGRIN INTEGRIN, CELL ADHESION, GLYCOPROTEIN
1087	1atz	A	358	516	4.2e-12	0.23	0.81		VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN-BINDING, HEMOSTASIS, DINUCLOTIDE BINDING FOLD
1087	1auq		345	552	4.2e-54	0.16	0.16		AI DOMAIN OF VON WILLEBRAND FACTOR; CHAIN: NULL;	WILLEBRAND WILLEBRAND, BLOOD COAGULATION, PLATELET, GLYCOPROTEIN
1087	1ck4	A	361	545	1.4e-31	0.37	0.42		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
1087	1d2i	A	358	534	3.4e-14	0.10	0.55		INTEGRIN; CHAIN: A; COLLAGEN; CHAIN: B, C, D;	INTEGRIN INTEGRIN, COLLAGEN
1087	1d2i	A	361	534	1.4e-28	0.23	1.00		INTEGRIN; CHAIN: A; COLLAGEN; CHAIN: B, C, D;	INTEGRIN INTEGRIN, COLLAGEN
1087	1fhs	A	355	549	7e-51	0.29	0.49		IMMUNOGLOBULIN NMC4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
1087	1ido		361	542	4.2e-35	0.13	0.48		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL

744

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1087	1lf8	A	361	547	7e-32	-0.04	0.46		CD11A; 1LFA 5 CHAIN; A, B; 1LFA 6	ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
1087	1qc5	A	361	543	5.6e-30	0.33	0.77		ALPHA1 BETA1 INTEGRIN; CHAIN: A; ALPHA1 BETA1 INTEGRIN; CHAIN: B;	CELL ADHESION INTEGRIN, CELL ADHESION
1088	1ciu		6	675	1.4e-77			82.96	CYCLODEXTRIN GLYCOSYLTRANSFERASE; 1CIU 6 CHAIN: NULL; 1CIU 7	GLYCOSIDASE CGTASE; 1CIU 8 THERMOSTABLE 1CIU 14
1088	1e43	A	9	484	9.8e-17			79.59	ALPHA-AMYLASE; CHAIN: A;	HYDROLASE HYDROLASE, AMYLASE, FAMILY 13
1088	1gcy	A	1	397	5.6e-18			74.07	GLUCAN 1,4-ALPHA-MALTOTETRAHYDROLASE; CHAIN: A;	HYDROLASE BETA-ALPHA-BARREL, BETA SHEET
1088	1hxo	A	5	488	5.6e-43			66.82	ALPHA AMYLASE (PPA); CHAIN: A;	HYDROLASE ALPHA-AMYLASE, INHIBITOR, CARBOHYDRATE, PANCREAS
1088	1qho	A	12	671	4.2e-70			81.19	ALPHA-AMYLASE; CHAIN: A;	HYDROLASE "MALTOGENIC" ALPHA AMYLASE; AMYLASE, GLYCOSIDE HYDROLASE, STARCH DEGRADATION
1088	1uok		110	675	0			90.36	OLIGO-1,6-GLUCOSIDASE; CHAIN: NULL;	GLUCOSIDASE GLUCOSIDASE, SUGAR DEGRADATION, HYDROLASE, TIM-BARREL 2 GLYCOSIDASE, HYDROLASE
1088	1uok		7	543	0			74.70	OLIGO-1,6-GLUCOSIDASE; CHAIN: NULL;	GLUCOSIDASE GLUCOSIDASE, SUGAR DEGRADATION, HYDROLASE, TIM-BARREL 2



Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1090	1aab		675	754	1.4e-17	0.90	1.00		HIGH MOBILITY GROUP PROTEIN; 1AAB 5 CHAIN: NULL; 1AAB 6	GLYCOSIDASE, HYDROLASE
1090	1aab		676	754	1.7e-23	0.83	1.00		HIGH MOBILITY GROUP PROTEIN; 1AAB 5 CHAIN: NULL; 1AAB 6	DNA-BINDING HMG A DNA-BINDING HMG-BOX DOMAIN A OF RAT HMG1; 1AAB 8 HMG-BOX 1AAB 20
1090	1eg7	A	669	751	4.2e-25	0.51	1.00		NON HISTONE PROTEIN 6 A; CHAIN: A;	DNA BINDING PROTEIN HMG BOX, DNA BENDING, DNA RECOGNITION, CHROMATIN, NMR, DNA 2 BINDING PROTEIN
1090	1ekt	A	680	748	1.4e-14	0.29	1.00		HIGH MOBILITY GROUP 1 PROTEIN; CHAIN: A; DNA (5'-D(*CP*CP*(IDO) CHAIN: B; DNA (5'-CHAIN: C;	GENE REGULATION/DNA HMG-1, AMPHOTERIN, HEPARIN-BINDING PROTEIN P30; HIGH-MOBILITY GROUP DOMAIN, BENT DNA, PROTEIN-DRUG-DNA 2 COMPLEX, GENE REGULATION/DNA
1090	1ekt	A	681	748	3.4e-20	0.47	1.00		HIGH MOBILITY GROUP 1 PROTEIN; CHAIN: A; DNA (5'-D(*CP*CP*(IDO) CHAIN: B; DNA (5'-CHAIN: C;	GENE REGULATION/DNA HMG-1, AMPHOTERIN, HEPARIN-BINDING PROTEIN P30; HIGH-MOBILITY GROUP DOMAIN, BENT DNA, PROTEIN-DRUG-DNA 2 COMPLEX, GENE REGULATION/DNA
1090	1hme		676	751	5.6e-29	0.50	1.00		DNA-BINDING HIGH MOBILITY GROUP PROTEIN FRAGMENT-B (HMG B) (DNA-BINDING 1HME 3 HMG-BOX DOMAIN B OF RAT HMG1) (NMR, 1 STRUCTURE) 1HME 4	
1090	1hsm		679	752	2.8e-27	0.68	1.00		DNA-BINDING HIGH	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									MOBILITY GROUP PROTEIN 1 (HMG1) BOX 2, COMPLEXED WITH IHSM 3 MERCAPTOETHANOL (NMR, MINIMIZED AVERAGE STRUCTURE) IHSM 4	
1090	1qrv	A	678	752	1.4e-17	0.65	1.00		DNA (5'- D(*G*CP*GP*AP*TP*AP *TP*CP*GP*C)-3'); CHAIN: C, D; HIGH MOBILITY GROUP PROTEIN D; CHAIN: A, B;	GENE REGULATION/DNA HMG-D; PROTEIN-DNA COMPLEX, HMG DOMAIN, NON-SEQUENCE SPECIFIC 2 CHROMOSOMAL PROTEIN HMG-D
1091	1d5s	B	342	382	5.6e-13			61.26	P1-ARG ANTITRYPSIN; CHAIN: A; P1-ARG ANTITRYPSIN; CHAIN: B;	HYDROLASE INHIBITOR SERPIN FOLD, RCL CLEAVAGE, A BETA SHEET POLYMERISATION
1091	1d5s	B	646	686	1.2e-14	-0.81	0.75		P1-ARG ANTITRYPSIN; CHAIN: A; P1-ARG ANTITRYPSIN; CHAIN: B;	HYDROLASE INHIBITOR SERPIN FOLD, RCL CLEAVAGE, A BETA SHEET POLYMERISATION
1091	1d5s	B	646	686	9.8e-13	-0.81	0.75		P1-ARG ANTITRYPSIN; CHAIN: A; P1-ARG ANTITRYPSIN; CHAIN: B;	HYDROLASE INHIBITOR SERPIN FOLD, RCL CLEAVAGE, A BETA SHEET POLYMERISATION
1091	1ezx	A	12	346	0			370.31	ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1- ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1- ANTITRYPSIN, 2 TRYPSIN
1091	1ezx	A	316	650	0			366.89	ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1- ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1- ANTITRYPSIN, 2 TRYPSIN

747

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1091	1ezx	A	317	650	0	0.69	1.00		ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1-ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
1091	1ezx	B	651	686	1.1e-11	-0.78	0.30		ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1-ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
1091	1ezx	B	651	686	3.4e-12	-0.78	0.30		ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1-ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
1091	1qlp	A	11	382	0			424.42	ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE, SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
1091	1qlp	A	315	686	0			424.95	ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE, SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
1091	1qlp	A	317	686	0	0.82	1.00		ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE, SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE

748

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
1091	1qmb	B	341	382	1.4e-12			61.02	ALPHA-1-ANTITRYPSIN; CHAIN: A, B;	SERINE PROTEASE INHIBITOR
									ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-Pi; SERPIN, ANTITRYPSIN, POLYMER, CLEAVED	
1091	1qmb	B	645	686	2.8e-12	-0.81	0.90		ALPHA-1-ANTITRYPSIN; CHAIN: A, B;	SERINE PROTEASE INHIBITOR
									ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-Pi; SERPIN, ANTITRYPSIN, POLYMER, CLEAVED	
1091	1qmb	B	645	686	5.1e-14	-0.81	0.90		ALPHA-1-ANTITRYPSIN; CHAIN: A, B;	SERINE PROTEASE INHIBITOR
									ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-Pi; SERPIN, ANTITRYPSIN, POLYMER, CLEAVED	
1092	1a0j	A	330	536	1.4e-69	0.19	0.65		TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
1092	1a0l	A	330	530	1.4e-67	0.07	0.76		BETA-TRYPTASE; CHAIN: A, B, C, D;	SERINE PROTEINASE TRYPSIN-LIKE SERINE PROTEINASE, TETRAMER, HEPARIN, ALLERGY, 2 ASTHMA
1092	1bru	P	330	536	2.8e-69	0.21	0.49		ELASTASE; CHAIN: P;	SERINE PROTEASE PPE, SERINE PROTEASE, HYDROLASE
1092	1ddj	A	328	536	7e-70	0.32	0.93		PLASMINOGEN; CHAIN: A, B, C, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC DOMAIN
1092	1dle	A	339	430	2.8e-17	0.39	-0.09		COMPLEMENT FACTOR B; CHAIN: A, B;	HYDROLASE SERINE PROTEASE, COMPLEMENT SYSTEM, FACTOR B, PROTEIN-2 PROTEIN

749

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1092	1dle	A	476	534	1.3e-07	-0.21	0.21		COMPLEMENT FACTOR B; CHAIN: A, B;	INTERACTION, ACTIVATION MECHANISM, BETA-BARREL FOLD,
1092	1elv	A	352	497	0.0034	0.17	0.89		COMPLEMENT C1S COMPONENT; CHAIN: A;	HYDROLASE TRYPSIN-LIKE SERIN PROTEASE, CCP (OR SUSHI OR SCR)MODULE
1092	1f7z	A	330	536	1.4e-67	0.34	0.95		TRYPSIN II, ANIONIC; CHAIN: A; PANCREATIC TRYPSIN INHIBITOR; CHAIN: I;	HYDROLASE/HYDROLASE INHIBITOR BPTI SERINE PROTEASE, TRYPSIN PRECURSOR
1092	1fh8	A	331	497	0.0017	0.66	0.84		TRYPSIN; CHAIN: A; GLY-ALA-ARG; CHAIN: B;	HYDROLASE BETA BARREL
1092	1fmi	A	330	536	1.3e-71	0.09	0.77		TRYPSIN; CHAIN: A;	HYDROLASE SERINE PROTEASE, HYDROLASE
1092	1qtf	A	357	532	3.4e-09	0.41	0.19		EXFOLIATIVE TOXIN B; CHAIN: A;	HYDROLASE, TOXIN EPIDERMOLYTIC TOXIN B; SERINE PROTEASE, SUPERANTIGEN, HYDROLASE, TOXIN
1092	1slw	B	330	536	5.6e-67	0.29	0.60		ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1092	1trn	A	330	536	1.4e-70	0.30	0.43		HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR ITRN 3 DISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) ITRN 4 HUMAN TRYPSIN, DFP INHIBITED ITRN 6	
1092	2sfa		357	532	1.7e-13	0.47	0.62		SERINE PROTEINASE; CHAIN: NULL;	HYDROLASE HYDROLASE, SERINE PROTEASE
1092	2sta	E	330	534	1.4e-68	0.17	0.76		TRYPSIN; CHAIN: E; TRYPSIN INHIBITOR; CHAIN: I	HYDROLASE/HYDROLASE INHIBITOR SERINE PROTEINASE, TRYPSIN INHIBITOR
1092	5ptp		330	536	2.8e-65	0.09	0.51		BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN SIGNAL HYDROLASE ALL ALPHA
1103	1d2t	A	27	242	3.4e-39	0.45	0.78		ACID PHOSPHATASE; CHAIN: A;	
1104	1cdq		36	112	2.8e-21			142.48	COMPLEMENT REGULATORY PROTEIN CD59 (NMR, 20 STRUCTURES) 1CDQ 3	
1104	1cdy		59	166	1e-08	0.23	-0.14		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
1104	1erg		36	105	1.4e-20			132.21	COMPLEMENT FACTOR HUMAN COMPLEMENT REGULATORY PROTEIN	

751

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1104	1f97	A	1	89	1.4e-33			54.30	CD59 (EXTRACELLULAR IERG 3 REGION, RESIDUES 1 - 70) (NMR, RESTRAINED MINIMIZED IERG 4 AVERAGE STRUCTURE) IERG 5	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1104	1f97	A	65	274	8.4e-52			269.92	JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1104	1f97	A	67	272	8.4e-52	0.92	1.00		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1104	1f97	A	67	274	3.4e-51	0.90	1.00		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1104	1wio	A	75	312	6.8e-28	0.01	-0.18		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
1105	1f5j	A	20	86	1.4e-29			87.22	APOLIPOPROTEIN CII; CHAIN: A;	LIPID TRANSPORT APOC-II; PROTEIN-LIPID INTERACTION, AMPHIPATHIC ALPHA HELIX
1105	1f5j	A	71	137	1.2e-30	-0.93	0.77		APOLIPOPROTEIN CII; CHAIN: A;	LIPID TRANSPORT APOC-II; PROTEIN-LIPID INTERACTION, AMPHIPATHIC ALPHA HELIX

752

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1105	1i5j	A	71	137	1.2e-30			86.49	APOLIPROTEIN CII; CHAIN: A;	LIPID TRANSPORT APOC-II; PROTEIN-LIPID INTERACTION, AMPHIPATHIC ALPHA HELIX
1105	1i5j	A	71	137	7e-29	-0.93	0.77		APOLIPROTEIN CII; CHAIN: A;	LIPID TRANSPORT APOC-II; PROTEIN-LIPID INTERACTION, AMPHIPATHIC ALPHA HELIX
1106	1h6q	A	1	159	2.8e-60	0.40	1.00		TRANSLATIONALLY CONTROLLED TUMOR PROTEIN; CHAIN: A;	TUMOR-ASSOCIATED PROTEIN TCTP, P23FYF; TUMOR-ASSOCIATED PROTEIN, FUNCTION UNKNOWN
1106	1h6q	A	1	160	1.7e-58	0.37	1.00		TRANSLATIONALLY CONTROLLED TUMOR PROTEIN; CHAIN: A;	TUMOR-ASSOCIATED PROTEIN TCTP, P23FYF; TUMOR-ASSOCIATED PROTEIN, FUNCTION UNKNOWN
1107	1awc	B	300	393	0.00051	-0.29	0.29		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABRALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
1107	1blx	B	337	411	0.00034	-0.03	0.09		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
1107	1dcq	A	337	397	0.00051	0.15	0.64		PYK2-ASSOCIATED PROTEIN BETA; CHAIN: A;	METAL BINDING PROTEIN ZINC-BINDING MODULE, ANKYRIN REPEATS, METAL BINDING PROTEIN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1107	1ika	D	337	394	0.00085	-0.47	0.55		NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFkB COMPLEX
1107	1myo		337	394	0.00068	0.18	0.11		MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
1107	1nfi	E	345	398	0.001	-0.25	0.84		NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
1107	1yes	B	337	423	0.00017	-0.05	0.03		P53; CHAIN: A; 53BP2; CHAIN: B;	COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS) P53BP2; ANKYRIN REPEATS, SH3, P53, TUMOR SUPPRESSOR, MULTIGENE 2 FAMILY, NUCLEAR PROTEIN, PHOSPHORYLATION, DISEASE MUTATION, 3 POLYMORPHISM, COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS)
1108	1d2h	A	108	240	1.4e-18	-0.26	0.06		GLYCINE N-METHYLTRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE METHYLTRANSFERASE
1109	1aj4		16	168	2.8e-26			99.38	TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC; CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING
1109	1aj4		97	222	2.8e-26	0.11	0.46		TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC; CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING

754

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1109	1ak8		3	74	5.6e-32			59.57	CALMODULIN; CHAIN: NULL;	CALCIUM-BINDING PROTEIN CALMODULIN CERIUM TRIC-DOMAIN, RESIDUES 1 - 75; CERIUM-LOADED, CALCIUM-BINDING PROTEIN
1109	1ap4		20	96	1.1e-18	0.79	1.00		CARDIAC N-TROPONIN C; CHAIN: NULL;	CALCIUM-BINDING CNTNC; CALCIUM-BINDING, REGULATION, TROPONIN C, CARDIAC MUSCLE 2 CONTRACTION
1109	1au1	B	18	179	2.8e-16			75.78	SERINE/THREONINE PHOSPHATASE 2B; CHAIN: A, B;	HYDROLASE CALCINEURIN; HYDROLASE, PHOSPHATASE, IMMUNOSUPPRESSION
1109	1avs	A	1	76	2.8e-28			54.68	TROPONIN C; CHAIN: A, B;	MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-ACTIVATED, TROPONIN, E-F HAND 2 CALCIUM-BINDING PROTEIN
1109	1blq		1	78	8.4e-29			50.01	N-TROPONIN C; CHAIN: NULL;	CALCIUM-BINDING PROTEIN SNTNC; CALCIUM-BINDING, REGULATION, TROPONIN C, SKELETAL MUSCLE, 2 CONTRACTION
1109	1br1	B	26	166	1.3e-38	0.63	1.00		MYOSIN; CHAIN: A, B, C; D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
1109	1br1	B	26	166	1.3e-38			92.46	MYOSIN; CHAIN: A, B, C; D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
1109	1br1	B	97	209	1.4e-11	0.24	0.22		MYOSIN; CHAIN: A, B, C; D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
1109	1cdm	A	102	209	1.3e-29	-0.08	0.19		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3	

755

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CALMODULIN-DEPENDENT PROTEIN KINASE II ICDM 4	
1109	1cdm	A	26	164	8.4e-59	0.72	1.00		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF ICDM 3	
									CALMODULIN-DEPENDENT PROTEIN KINASE II ICDM 4	
1109	1cdm	A	26	164	8.4e-59			118.25	CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF ICDM 3	
									CALMODULIN-DEPENDENT PROTEIN KINASE II ICDM 4	
1109	1cll		102	209	9.8e-36	-0.05	0.24		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
									CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
1109	1cll		15	91	1.4e-19	0.39	1.00		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
									CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
1109	1cll		1	86	1.4e-42			50.32	CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
									CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
1109	1cll		26	164	5.6e-65	0.73	1.00		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
									CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	
1109	1cll		26	165	5.6e-65			135.44	CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) ICLL 3	

756

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1109	1cmf		15	87	1.4e-05			70.55	CALMODULIN (VERTEBRATE); 1CMF 6 CHAIN: NULL; 1CMF 7	CALCIUM-BINDING PROTEIN CALMODULIN APO TR2C-DOMAIN; 1CMF 9
1109	1dgu	A	12	177	8.4e-16			64.07	CALCIUM-SATURATED CIB; CHAIN: A	BLOOD CLOTTING HELICAL, EF-HANDS, BLOOD CLOTTING
1109	1dtl	A	20	165	2.8e-26			91.37	CARDIAC TROPONIN C; CHAIN: A;	STRUCTURAL PROTEIN HELIX-TURN-HELIX
1109	1dtl	A	97	222	2.8e-26	0.36	0.63		CARDIAC TROPONIN C; CHAIN: A;	STRUCTURAL PROTEIN HELIX-TURN-HELIX
1109	1ext	A	102	209	2.8e-33	0.04	0.29		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
1109	1ext	A	15	90	4.2e-18	0.29	0.95		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
1109	1ext	A	24	163	1.4e-62	0.75	1.00		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
1109	1ext	A	24	165	1.4e-62			132.92	CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
1109	1f4q	A	3	130	3.4e-10	-0.15	0.30		GRANCALCIN; CHAIN: A, B;	METAL TRANSPORT PENTA-EF-HAND PROTEIN, CALCIUM BINDING PROTEIN
1109	1fpw	A	48	216	8.4e-19	-0.21	0.37		CALCIUM-BINDING PROTEIN NCS-1; CHAIN: A;	METAL BINDING PROTEIN YEAST FREQUENIN EF-HAND, CALCIUM
1109	1fw4	A	20	84	1.4e-05			67.19	CALMODULIN; CHAIN: A;	METAL BINDING PROTEIN EF-HAND, HELIX-LOOP-HELIX, FRAGMENT, CALCIUM, TR2C, C-2 TERMINAL DOMAIN, CALMODULIN
1109	1g8i	A	6	180	7e-14			65.19	NEURONAL CALCIUM SENSOR 1; CHAIN: A, B;	METAL BINDING PROTEIN FREQUENIN; CALCIUM BINDING-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										PROTEIN, EF-HAND, CALCIUM ION
1109	1ggw	A	28	166	7e-11			89.53	CDC4P; CHAIN: A;	CYTOKINE EF-HAND PROTEIN, MYOSIN LIGHT CHAIN; LIGHT CHAIN, CYTOKINESIS, CELL CYCLE, EF-HAND
1109	1hqy	A	15	203	2.8e-23			60.78	PROGRAMMED CELL DEATH PROTEIN 6; CHAIN: A;	APOPTOSIS PROBABLE CALCIUM-BINDING PROTEIN ALG-2; PENTA-EF-HAND PROTEIN, CALCIUM BINDING PROTEIN
1109	1hqy	A	37	194	2.8e-23	0.19	0.24		PROGRAMMED CELL DEATH PROTEIN 6; CHAIN: A;	APOPTOSIS PROBABLE CALCIUM-BINDING PROTEIN ALG-2; PENTA-EF-HAND PROTEIN, CALCIUM BINDING PROTEIN
1109	1lku		5	191	5.6e-11			58.57	RECOVERIN; CHAIN: NULL;	CALCIUM-BINDING PROTEIN CALCIUM-MYRISTOYL SWITCH, CALCIUM-BINDING PROTEIN
1109	1lcf		102	224	7e-30	0.06	0.40		TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION
1109	1lcf		17	165	7e-30			104.72	TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION
1109	1top		102	224	2.8e-30	0.15	0.87		CONTRACTILE SYSTEM	CONTRACTON MUSCLE BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION

758

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEIN TROPONIN C 1TOP 3	
1109	1top		13	168	2.8e-30			107.77	CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
1109	1trc	A	19	86	1.4e-05			63.97	CALCIUM BINDING PROTEIN CALMODULIN (1TR=2=C\$ FRAGMENT COMPRISING RESIDUES 78 - 148 1TRC 3 OF THE INTACT MOLECULE) 1TRC 4	
1109	1trf		5	76	2.8e-28			53.23	MUSCLE PROTEIN TROPONIN C (TRIC FRAGMENT) (APO FORM) (NMR, 1 STRUCTURE) 1TRF 3	
1109	1vfk	A	102	209	2.8e-34	0.24	0.57		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
1109	1vfk	A	15	93	1.1e-18	0.39	0.99		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
1109	1vfk	A	23	166	9.8e-64	0.60	1.00		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
1109	1vfk	A	24	166	9.8e-64			133.11	CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										SIGNALING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
1109	1wdc	B	1	89	7e-20			67.20	SCALLOP MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MYOSIN, CALCIUM BINDING PROTEIN, MUSCLE PROTEIN
1109	1wdc	B	26	168	1.7e-43			163.19	SCALLOP MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MYOSIN, CALCIUM BINDING PROTEIN, MUSCLE PROTEIN
1109	1wdc	B	26	168	5.6e-35	0.48	1.00		SCALLOP MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MYOSIN, CALCIUM BINDING PROTEIN, MUSCLE PROTEIN
1109	1wdc	B	28	166	1.7e-43	0.52	1.00		SCALLOP MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MYOSIN, CALCIUM BINDING PROTEIN, MUSCLE PROTEIN
1109	1wdc	C	26	169	1.3e-06			89.97	SCALLOP MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MYOSIN, CALCIUM BINDING PROTEIN, MUSCLE PROTEIN
1109	2mys	B	1	90	2.8e-18			51.56	MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
1109	2mys	B	26	166	1.7e-38	-0.25	1.00		MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
1109	2mys	B	26	169	1.7e-38			146.55	MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
1109	2mys	B	7	78	5.6e-22			50.77	MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN

760

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1109	2mys	B	96	198	5.6e-17	-0.25	0.05		MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
1109	2mys	C	29	165	2.8e-35			87.97	MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
1109	2mys	C	32	165	2.8e-35	-0.05	1.00		MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
1112	1ajj		116	151	5.1e-09	-0.28	0.13		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: NULL;	RECEPTOR LR5; RECEPTOR, LDL RECEPTOR, CYSTEINE-RICH MODULE, CALCIUM
1112	1ajj		117	151	5.6e-09	-0.06	0.01		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: NULL;	RECEPTOR LR5; RECEPTOR, LDL RECEPTOR, CYSTEINE-RICH MODULE, CALCIUM
1112	1f8z	A	117	151	1.3e-07	-0.12	0.06		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A;	LIPID BINDING PROTEIN LDL RECEPTOR, LIGAND-BINDING DOMAIN, CALCIUM-BINDING, 2 FAMILIAL
1112	1ldl		116	151	5.1e-07	0.29	0.33		LOW-DENSITY LIPOPROTEIN RECEPTOR; 1LDL 4 CHAIN; NULL; 1LDL 5	BINDING PROTEIN LB1; 1LDL 7 LDL RECEPTOR CYSTEINE-RICH REPEAT 1LDL 15
1112	1sfp		1	113	8.4e-07	0.35	0.04		ASFP; CHAIN: NULL;	SPERMADHESIN ACIDIC SEMINAL PROTEIN; SPERMADHESIN, BOVINE SEMINAL PLASMA PROTEIN, ACIDIC 2 SEMINAL FLUID



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1112	Isfp		26	114	1.7e-10	0.37	0.09		ASFP; CHAIN: NULL;	PROTEIN, ASFP, CUB DOMAIN, X-RAY CRYSTAL 3 STRUCTURE, GROWTH FACTOR
										SPERMADHESIN ACIDIC SEMINAL PROTEIN; SPERMADHESIN, BOVINE SEMINAL PLASMA PROTEIN, ACIDIC 2 SEMINAL FLUID PROTEIN, ASFP, CUB DOMAIN, X-RAY CRYSTAL 3 STRUCTURE, GROWTH FACTOR
1112	Ispp	A	26	112	8.5e-09	0.35	0.30		MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	COMPLEX (SEMINAL PLASMA PROTEIN/SP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN/SP)
1112	Ispp	B	26	112	5.1e-10	0.17	0.11		MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	COMPLEX (SEMINAL PLASMA PROTEIN/SP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN/SP)
1119	leis	A	277	354	9.8e-07	0.46	-0.09		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
1119	9wga	A	485	655	4.2e-11	0.19	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1119	9wga	A	915	1107	1.4e-13	0.04	-0.19		LECTIN (AGGLUTININ) WHEAT GERM	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1123	1e08	A	1	325	0			152.98	[FE]-HYDROGENASE (LARGE SUBUNIT); CHAIN: A; [FE]-HYDROGENASE (SMALL SUBUNIT); CHAIN: D; CYTOCHROME C553; CHAIN: E	HYDROGENASE HYDROGENASE, CYTOCHROME C553, ELECTRON TRANSFER COMPLEX
1123	1hfe	L	1	325	0			137.95	FE-ONLY HYDROGENASE (SMALLER SUBUNIT); CHAIN: S, T; FE-ONLY HYDROGENASE (LARGER SUBUNIT); CHAIN: I, M;	HYDROGENASE FE-ONLY HYDROGENASE, X-RAY CRYSTALLOGRAPHY, HYDROGENE 2 METABOLISM, PERIPLASM
1123	1j5j	A	8	217	0.0068			51.59	SENSORY RHODOP SIN II; CHAIN: A;	SIGNALING PROTEIN SENSORY RHODOP SIN, MEMBRANE PROTEIN, PHOTOTAXIS RECEPTOR
1124	1dv8	A	199	326	5.6e-34	-0.18	0.33		ASIALOGLYCOPROTEIN RECEPTOR I; CHAIN: A;	SIGNALING PROTEIN HEPATIC LECTIN HI; C-TYPE LECTIN CRD
1124	1hq8	A	194	308	5.6e-28			86.06	NKG2-D; CHAIN: A;	APOPTOSIS HOMODIMER, CIS-PROLINE
1124	1h9r	A	193	315	2.8e-26			100.53	NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-RELATED PROTEIN A; CHAIN: C;	IMMUNE SYSTEM NKG2D; MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A, MHC-I, COMPLEX, IMMUNE SYSTEM
1124	1h9r	A	48	163	1.4e-27			94.69	NKG2-D TYPE II INTEGRAL MEMBRANE PROTEIN; CHAIN: B, A; MHC CLASS I CHAIN-	IMMUNE SYSTEM NKG2D; MIC-A, MIC, PERB11; ACTIVATING NK CELL RECEPTOR, NKG2D, C-TYPE-LECTIN LIKE, MIC-2 A,

763

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									RELATED PROTEIN A; CHAIN: C;	MHC-I, COMPLEX, IMMUNE SYSTEM
1125	12e8	L	71	269	9.8e-17			73.24	2E8 (IGG1=KAPPA=) ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN IMMUNOGLOBULIN
1125	12e8	L	83	267	9.8e-17	0.28	0.75		2E8 (IGG1=KAPPA=) ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN IMMUNOGLOBULIN
1125	1adq	L	72	258	2.8e-22	0.38	1.00		IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
1125	1adq	L	72	271	2.8e-22			72.32	IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
1125	1b2w	L	70	269	4.2e-20			73.56	ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
1125	1b6d	A	70	269	7e-21			73.38	IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
1125	1b1h	A	2	362	1.3e-43	0.06	0.99		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING,

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1125	1b1h	A	2	364	1.3e-43			121.44	HEMOLIN; CHAIN: A, B;	HOMOPHILIC ADHESION
										INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1125	1b1h	A	73	386	5.1e-38	0.18	0.77		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1125	1b1l	J	76	267	1.4e-21	0.10	0.89		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
1125	1b1l	H	84	268	9.8e-14	0.01	0.34		COMPLEX (ANTIBODY/ANTIGEN) HYHEL-5 FAB COMPLEXED WITH BOBWHITE QUAIL LYSOZYME 1BQL 3 1BQL 95	
1125	1bz7	A	70	265	4.2e-19			73.01	ANTIBODY R24 (LIGHT CHAIN); CHAIN: A; ANTIBODY R24 (HEAVY CHAIN); CHAIN: B;	IMMUNE SYSTEM ANTIBODY (FAB FRAGMENT), IMMUNE SYSTEM
1125	1c1c	A	83	267	5.6e-17	0.35	0.82		IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D; AXONIN-1; CHAIN: A;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
1125	1cs6	A	10	363	5.6e-41	0.10	0.86			CELL ADHESION NEURAL CELL ADHESION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1125	1cs6	A	15	364	1e-43			99.18	AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1125	1cs6	A	66	429	5.6e-40	0.19	0.76		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1125	1cs6	A	72	379	1e-43	0.09	0.16		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1125	1cvs	C	169	362	4.2e-38	0.15	0.48		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1125	1cvs	C	6	167	8.4e-20	0.01	-0.09		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1125	1cvs	C	81	270	4.2e-22	0.02	-0.05		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1125	1cvs	D	169	362	1.3e-39	0.32	0.80		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR

766

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1125	1cvs	D	6	167	2.8e-20	-0.26	0.06		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1125	1d5i	L	70	269	4.2e-21			72.82	CHIMERIC GERM LINE PRECURSOR OF OXY-COPE CHAIN: L; CHIMERIC GERM LINE PRECURSOR OF OXY-COPE CHAIN: H;	IMMUNE SYSTEM IMMUNE SYSTEM
1125	1dfb	L	70	269	8.4e-22			75.44	IMMUNOGLOBULIN 3D6	
1125	1dfb	L	76	267	8.4e-22	0.36	0.99		IMMUNOGLOBULIN 3D6	
1125	1dgi	R	58	362	3.4e-51			111.26	FAB 1DFB 3	
									POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1125	1dgi	R	75	362	3.4e-51	-0.21	0.46		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1125	1dgi	R	76	362	1.4e-39	-0.02	0.06		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1125	1epf	A	165	352	1.7e-28	0.36	0.55		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1125	1epf	A	175	346	2.8e-19	0.37	0.94		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1125	1epf	A	3	152	2.8e-15	-0.07	0.00		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1125	1epf	A	72	272	4.2e-24	-0.00	0.16		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1125	1ev2	E	170	362	2.8e-34	0.04	0.23		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOIL FOLD
1125	1ev2	G	170	366	4.2e-37	0.30	0.70		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOIL FOLD
1125	1evt	C	169	362	2.8e-39	0.03	0.51		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOIL FOLD
1125	1t2g	A	26	171	2.8e-12	-0.13	0.07		HIGH AFFINITY	IMMUNE SYSTEM FC-EPSILON RL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1125	1f6a	A	166	365	3.4e-27	0.26	0.18		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IGE EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1125	1f6a	A	171	346	1.4e-14	0.40	0.99		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IGE EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1125	1f97	A	181	362	5.1e-26	0.27	0.11		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1125	1f97	A	5	158	2.8e-12	-0.05	0.09		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1125	1f97	A	77	265	4.2e-30	0.09	0.86		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1125	1f6g	A	170	362	8.5e-28	0.14	0.39		FC RECEPTOR FC(GAMMA)RIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32, FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1125	1f6g	A	272	362	1.5e-17	0.48	0.74		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1125	1fhg	A	275	362	2.8e-17	0.38	0.72		TELOKIN; CHAIN: A	BARREL CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
1125	1fhg	A	78	167	5.6e-13	0.02	0.33		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
1125	1fml	A	167	362	3.4e-26	0.21	0.01		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1125	1fml	A	273	375	1.7e-16	0.34	-0.06		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1125	1g0x	A	167	356	6.8e-24	0.23	0.05		LEUCOCYTE IMMUNOGLOBULIN- LIKE RECEPTOR-1; CHAIN: A;	IMMUNE SYSTEM LEUCOCYTE INHIBITORY RECEPTOR-1; LEUCOCYTE IMMUNOGLOBULIN FOLD, 3-10 HELIX
1125	1iai	L	83	267	9.8e-15	0.12	0.27		IDIOTYPIC FAB 730.1.4 (IGG1) OF VIRUS IIAI 5 CHAIN: L, H; IIAI 7 ANTI-IDIOTYPIC FAB 409.5.3 (GG2A); IIAI 9 CHAIN: M, I IIAI 10	COMPLEX (IMMUNOGLOBULIN IGG1/IG2A)
1125	1ie5	A	269	362	8.4e-18	-0.11	0.45		NEURAL CELL ADHESION MOLECULE; CHAIN: A;	CELL ADHESION N-CAM; INTERMEDIATE IMMUNOGLOBULIN FOLD
1125	1ie5	A	272	363	6.8e-18	0.01	0.51		NEURAL CELL ADHESION MOLECULE; CHAIN: A;	CELL ADHESION N-CAM; INTERMEDIATE IMMUNOGLOBULIN FOLD
1125	1iil	G	164	366	6.8e-27	0.24	0.46		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF- 2, BASIC FIBROBLAST GROWTH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-1REFOL.
1125	1iil	G	170	366	1.4e-36	0.38	0.53		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-1REFOL.
1125	1iil	G	275	365	1.5e-16	0.61	0.55		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-1REFOL.
1125	1iib	B	79	366	3.4e-37			82.33	INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1125	1mco	H	1	363	9.8e-19			78.42	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1125	1nfd	E	74	267	1.1e-21	0.32	0.96		N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1125	1osp	L	70	269	1.1e-17			75.72	FAB 184.1; CHAIN: L, H; OUTER SURFACE PROTEIN A; CHAIN: O;	COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN) OSP A; COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN), OUTER SURFACE 2 PROTEIN A COMPLEXED WITH FAB184.1, BORRELLIA BURGDORFERI 3 STRAIN B31
1125	1vca	A	73	278	8.5e-27	0.57	0.93		HUMAN VASCULAR CELL ADHESION MOLECULE-1; 1VCA 4 CHAIN: A, B; 1VCA 5	CELL ADHESION PROTEIN VCAM-D1,2; 1VCA 6 IMMUNOGLOBULIN SUPERFAMILY, INTEGRIN-BINDING 1VCA 15
1125	1wio	A	76	442	1.5e-35			89.83	T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
1125	2dli	A	167	354	8.5e-23	0.17	0.00		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR, KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
1125	2fcb	A	170	365	1.5e-27	-0.16	0.75		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1125	2fcb	A	278	375	1.5e-16	0.17	0.03		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1125	2fgw	L	76	267	1.1e-21	0.30	0.99		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	

772

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1125	2ncm		282	363	1.5e-17	0.42	0.29		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
1125	3fct	A	73	269	1.4e-20			77.03	METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: A, C; METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: B, D;	IMMUNE SYSTEM METAL CHELATASE, CATALYTIC ANTIBODY, FAB FRAGMENT, IMMUNE 2 SYSTEM
1125	8fab	A	73	268	5.6e-23			74.31	IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
1125	8fab	A	75	258	5.6e-23	0.42	1.00		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
1126	1b3u	A	22	571	4.5e-18	0.11	-1202.08		PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT REPEAT
1126	1ee4	A	389	777	4.5e-21	0.35	-1202.08		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
1126	1g3j	C	462	799	3e-15	0.02	-1202.08		BETA-CATENIN	TRANSCRIPTION BETA-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ARMADILLO REPEAT REGION; CHAIN: A, C, TCF3-CBD (CATENIN BINDING DOMAIN); CHAIN: B, D;	CATENIN,TCF-3, PROTEIN-PROTEIN COMPLEX
1126	1i7w	A	462	915	1.5e-21	0.09	-1202.08		BETA-CATENIN; CHAIN: A, C; EPITHELIAL-CADHERIN; CHAIN: B, D;	CELL ADHESION E-CADHERIN; E-CADHERIN, CELL ADHESION, BETA-CATENIN, PROTEIN-PROTEIN 2 COMPLEX, EXTENDED INTERFACE, ARMADILLO REPEAT, PHOSPHOSERINE
1126	1ial	A	456	901	1.5e-18	0.14	-1202.08		IMPORTIN ALPHA; CHAIN: A;	NUCLEAR IMPORT RECEPTOR KARYOPHERIN ALPHA; NUCLEAR IMPORT RECEPTOR, NUCLEAR LOCALIZATION SIGNAL, 2 ARMADILLO REPEATS, AUTOINHIBITION, INTRASTERIC REGULATION
1126	3bct		412	787	6e-17	0.11	-1202.08		BETA-CATENIN; CHAIN: NULL;	ARMADILLO REPEAT ARMADILLO REPEAT, BETA-CATENIN, CYTOSKELETON
1127	1a7q	L	27	132	0.00012			60.18	MONOCLONAL ANTIBODY D1.3; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, VARIANT
1127	1aif	A	27	211	0.0015			61.80	ANTI-IDIOTYPIC FAB 409.5.3 (GG2A) FAB; CHAIN: A, B, L, H	IMMUNOGLOBULIN IMMUNOGLOBULIN, C REGION, V REGION
1127	1bwv	A	25	133	0.00045			61.39	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENGE-JONES 2 PROTEIN, IMMUNE SYSTEM
1127	1cdy		35	136	1.5e-09	0.41	-1202.08		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD,

Table 5

SEQ ID NO:	PDB ID	CHAIN NID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
1127	1cs6	A	26	159	6e-08	0.25	-1202.08		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1127	1cvs	C	9	112	1.5e-11	0.04	-1202.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1127	1dr9	A	37	134	4.5e-08	0.35	-1202.08		T LYMPHOCYTE ACTIVATION ANTIGEN; CHAIN: A;	IMMUNE SYSTEM B7-1 (CD80); IG SUPERFAMILY
1127	1eaj	A	28	132	7.5e-11	0.41	-1202.08		COXSACKIE VIRUS AND ADENOVIRUS RECEPTOR; CHAIN: A, B;	VIRUS/VIRAL PROTEIN RECEPTOR COXSACKIEVIRUS B-ADENOVIRUS RECEPTOR, HCAR, VIRUS/VIRAL PROTEIN RECEPTOR, IMMUNOGLOBULIN V DOMAIN FOLD, 2 SYMMETRIC DIMER
1127	1epf	A	31	112	6e-10	0.47	-1202.08		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1127	1f97	A	30	112	1.5e-10	0.12	-1202.08		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1127	1g9m	L	27	210	3e-06			61.27	ENVELOPE GLYCOPROTEIN GP120; CHAIN: G; T-CELL SURFACE	VIRUS/VIRAL PROTEIN COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB), HIV-1 EXTERIOR 2 ENVELOPE GP120

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GLYCOPROTEIN CD4; CHAIN: C; ANTIBODY 17B, LIGHT CHAIN; CHAIN: L; ANTIBODY 17B, HEAVY CHAIN; CHAIN: H;	FROM LABORATORY-ADAPTED ISOLATE, HXBC2, 3 SURFACE T-CELL GLYCOPROTEIN CD4, ANTIGEN-BINDING FRAGMENT 4 OF HUMAN IMMUNOGLOBULIN 17B
1127	1hxm	B	32	142	6e-10	0.32	-1202.08		GAMMA-DELTA T-CELL RECEPTOR; CHAIN: A, C, E, G; GAMMA-DELTA T-CELL RECEPTOR; CHAIN: B, D, F, H;	IMMUNE SYSTEM T-CELL RECEPTOR DELTA CHAIN; T-CELL RECEPTOR GAMMA CHAIN; IG DOMAIN, T CELL RECEPTOR, TCR, GDTOR
1127	1igm	L	27	140	7.5e-05			60.22	IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
1127	1neu		31	132	1.2e-09	0.36	-1202.08		MYELIN P0 PROTEIN; CHAIN: NULL;	STRUCTURAL PROTEIN MYELIN, STRUCTURAL PROTEIN, GLYCOPROTEIN, TRANSMEMBRANE, PHOSPHORYLATION, IMMUNOGLOBULIN FOLD, SIGNAL, MYELIN 2 MEMBRANE ADHESION MOLECULE
1127	1nkr		29	148	1.5e-09	0.16	-1202.08		P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
1127	1vca	A	31	134	1.5e-10	0.22	-1202.08		HUMAN VASCULAR CELL ADHESION MOLECULE-1; 1VCA 4 CHAIN: A, B; 1VCA 5	CELL ADHESION PROTEIN VCAM-D12; 1VCA 6 IMMUNOGLOBULIN SUPERFAMILY, INTEGRIN-BINDING 1VCA 15
1131	1a25	A	648	768	1.4e-17	0.24	-1202.08		PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALB;

776

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1131	1a25	A	667	754	1.5e-18	0.08	-1202.08		PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM+/+PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN
1131	1byp	A	648	768	8.4e-23	0.21	-1202.08		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1131	1c1y	A	666	786	1.4e-08	0.16	-1202.08		CYTOSOLIC PHOSPHOLIPASE A2; CHAIN: A, B;	HYDROLASE CPLA2; PHOSPHOLIPASE, LIPID-BINDING, HYDROLASE
1131	1d1x	B	845	1047	2.8e-12	0.01	-1202.08		PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C, CHAIN: A, B;	LIPID DEGRADATION PLC-D1; PHOSPHORIC DIESTER HYDROLASE, HYDROLASE, LIPID DEGRADATION, 2 TRANSDUCER, CALCIUM-BINDING, PHOSPHOLIPASE C, 3
1131	1d5y	A	647	781	5.6e-21	0.23	-1202.08		PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	TRANSFERASE CALCIUM+, PHOSPHOLIPID BINDING PROTEIN, CALCIUM-BINDING 2 PROTEIN, PHOSPHATIDYL SERINE, PROTEIN KINASE C
1131	1d1w		664	754	4.5e-19	0.04	-1202.08		PHOSPHOLIPASE A2; CHAIN: NUCL;	HYDROLASE CALB DOMAIN; HYDROLASE, C2 DOMAIN, CALB DOMAIN
1131	1d1w		666	765	2.8e-08	0.14	-1202.08		PHOSPHOLIPASE A2;	HYDROLASE CALB DOMAIN;



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1131	1rsy		619	754	3e-20	0.22	-1202.08		CHAIN: NULL;	HYDROLASE, C2 DOMAIN, CALB DOMAIN
1131	1rsy		648	768	8.4e-23	0.11	-1202.08		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN 1 (FIRST C2 DOMAIN) (CALB) IRSY 3	
1131	3rpb	A	650	779	9.8e-17	0.29	-1202.08		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN 1 (FIRST C2 DOMAIN) (CALB) IRSY 3	
1133	1f88	A	54	378	1e-24			73.00	RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
1133	1f88	B	54	371	3e-18			70.57	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1133	1hme		6	81	1.1e-28			98.43	DNA-BINDING HIGH MOBILITY GROUP PROTEIN FRAGMENT-B (HMGB) (DNA-BINDING 1HME 3 HMGB-BOX DOMAIN B OF RAT HMGI) (NMR, 1 STRUCTURE) 1HME 4	PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1133	1hsm		9	87	1.1e-26			97.62	DNA-BINDING HIGH MOBILITY GROUP PROTEIN 1 (HMG1) BOX 2, COMPLEXED WITH IHSM3	
									MERCAPTOETHANOL (NMR, MINIMIZED AVERAGE STRUCTURE) IHSM4	
1134	1fx8	A	85	333	1.4e-47			73.96	GLYCEROL UPTAKE FACILITATOR PROTEIN; CHAIN: A;	MEMBRANE PROTEIN GLPF; GLYCEROL-CONDUCTING MEMBRANE CHANNEL PROTEIN
1138	1914		1	104	9.8e-36			69.38	SIGNAL RECOGNITION PARTICLE 9/14 FUSION PROTEIN; CHAIN: NULL;	ALU DOMAIN SRP9/14, ALU BM, RBD; ALU DOMAIN, CRYSTAL STRUCTURE, RNA BINDING, SIGNAL 2 RECOGNITION PARTICLE (SRP), TRANSLATION REGULATION
1138	1dhp	A	35	327	1.4e-91	0.66	-1202.08		DIHYDRODIPICOLINATE SYNTHASE; CHAIN: A, B;	SYNTHASE DHDP5; SYNTHASE, DIHYDRODIPICOLINATE
1138	1dhp	A	35	327	1.4e-91			130.43	DIHYDRODIPICOLINATE SYNTHASE; CHAIN: A, B;	SYNTHASE DHDP5; SYNTHASE, DIHYDRODIPICOLINATE
1138	1f6k	A	10	304	1.4e-69			123.19	N-ACETYLNEURAMINATE LYASE; CHAIN: A, C;	LYASE BETA BARREL, LYASE
1138	1f6k	A	33	327	3e-76			123.09	N-ACETYLNEURAMINATE LYASE; CHAIN: A, C;	LYASE BETA BARREL, LYASE
1138	1f6k	A	34	318	3e-76	0.52	-1202.08		N-ACETYLNEURAMINATE LYASE; CHAIN: A, C;	LYASE BETA BARREL, LYASE
1138	1f6k	A	34	323	5.6e-69	0.44	-1202.08		N-ACETYLNEURAMINATE LYASE; CHAIN: A, C;	LYASE BETA BARREL, LYASE

779

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ACETYLNEURAMINATE LYASE; CHAIN: A, C;	
1138	Inal	1	11	300	7e-64			121.57	N-ACETYLNEURAMINATE LYASE; INAL 4 CHAIN: 1, 2, 3, 4; INAL 5	LYASE
1138	Inal	1	34	318	4.5e-75	0.66	-1202.08		N-ACETYLNEURAMINATE LYASE; INAL 4 CHAIN: 1, 2, 3, 4; INAL 5	LYASE
1138	Inal	1	34	319	1.4e-63	0.61	-1202.08		N-ACETYLNEURAMINATE LYASE; INAL 4 CHAIN: 1, 2, 3, 4; INAL 5	LYASE
1138	Inal	1	34	323	4.5e-75			121.46	N-ACETYLNEURAMINATE LYASE; INAL 4 CHAIN: 1, 2, 3, 4; INAL 5	LYASE
1140	lhci	A	363	478	1.5e-08	0.04	-1202.08		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE 2 Z-LINE, ACTIN-BINDING PROTEIN
1142	lhx5	I	176	284	1.4e-10	0.44	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C; D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1142	1dx5	I	252	353	5.6e-14	0.33	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
1142	1dx5	I	320	427	5.6e-12	0.33	-1202.08		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
1142	1hj7	A	214	288	1.4e-09	0.33	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
1142	1hj7	A	368	427	1.3e-09	0.02	-1202.08		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
1142	1k1o		163	288	5.6e-09	0.18	-1202.08		LAMININ; CHAIN: NULL; FACTOR IXA; CHAIN: C, L; D-PHE-PRO-ARG; CHAIN: I;	GLYCOPROTEIN GLYCOPROTEIN COMPLEX (BLOOD COAGULATION/INHIBITOR)
1142	1pfx	L	250	327	8.4e-09	0.09	-1202.08			CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1142	9wga	A	290	462	7e-15	0.01	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLLECTIN 2) 9WGA 3	CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
1142	9wga	A	98	263	2.8e-15	-0.00	-1202.08		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLLECTIN 2) 9WGA 3	
1143	1f88	A	1	275	7e-82			58.80	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN- COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1143	1f88	A	25	366	1.4e-90			95.22	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN- COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1143	1f88	B	23	352	2.8e-82			66.80	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN- COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1145	1a9c	A	661	780	6e-23	0.03	-1202.08		X11; CHAIN: A, B; PEPTIDE; CHAIN: C, D	COMPLEX (PEPTIDE BINDING MODULE/PEPTIDE), PEPTIDE BINDING 2 MODULE, PTB DOMAIN
1145	1ddm	A	484	618	1.5e-13	0.11	-1202.08		NUMB PROTEIN; CHAIN: A; NUMB ASSOCIATE KINASE; CHAIN: B;	SIGNALING PROTEIN/TRANSFERASE NAK; COMPLEX, SIGNAL TRANSDUCTION, PHOSPHOTYROSINE BINDING 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1145	1ddm	A	661	778	4.5e-25	0.45	-1202.08		NUMB PROTEIN; CHAIN: A; NUMB ASSOCIATE KINASE; CHAIN: B;	SIGNALING PROTEIN/TRANSFERASE NAK; COMPLEX, SIGNAL TRANSDUCTION, PHOSPHOTYROSINE BINDING 2 DOMAIN (PTB), ASYMMETRIC CELL DIVISION
1145	1shc	A	477	620	9e-21	0.42	-1202.08		SHC; CHAIN: A; TRKA RECEPTOR PHOSHOPEPTIDE; CHAIN: B;	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), PHOSPHOTYROSINE 2 BINDING DOMAIN (PTB)
1145	1x11	A	661	782	3e-23	0.20	-1202.08		X11; CHAIN: A, B; 13-MER PEPTIDE; CHAIN: C, D;	COMPLEX (PEPTIDE BINDING MODULE/PEPTIDE), PTB DOMAIN
1145	2nmb	A	661	786	7.5e-26	0.28	-1202.08		NUMB PROTEIN; CHAIN: A; GPPY PEPTIDE; CHAIN: B;	CELL CYCLE/GENE REGULATION COMPLEX, SIGNAL TRANSDUCTION, PHOSPHOTYROSINE BINDING 2 DOMAIN (PTB), ASYMETRIC CELL DIVISION
1148	1a25	A	224	347	4.2e-24	0.06	-1202.08		PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALB; CALCIUM++/PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN
1148	1a25	A	72	191	4.2e-24	0.15	-1202.08		PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALB; CALCIUM++/PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1148	1bvm	A	69	185	8.4e-30	0.21	-1202.08		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1148	1dix	A	255	356	2.8e-21	0.32	-1202.08		PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C, CHAIN: A, B;	LIPID DEGRADATION PLC-D1; PHOSPHORIC DIESTER HYDROLASE, HYDROLASE, LIPID DEGRADATION, 2 TRANSDUCER, CALCIUM-BINDING, PHOSPHOLIPASE C, 3 PHOSPHOINOSITIDE-SPECIFIC
1148	1dix	B	115	355	1e-31	0.00	-1202.08		PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C, CHAIN: A, B;	LIPID DEGRADATION PLC-D1; PHOSPHORIC DIESTER HYDROLASE, HYDROLASE, LIPID DEGRADATION, 2 TRANSDUCER, CALCIUM-BINDING, PHOSPHOLIPASE C, 3 PHOSPHOINOSITIDE-SPECIFIC
1148	1dix	B	255	356	2.8e-21	0.38	-1202.08		PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C, CHAIN: A, B;	LIPID DEGRADATION PLC-D1; PHOSPHORIC DIESTER HYDROLASE, HYDROLASE, LIPID DEGRADATION, 2 TRANSDUCER, CALCIUM-BINDING, PHOSPHOLIPASE C, 3 PHOSPHOINOSITIDE-SPECIFIC
1148	1dsy	A	70	194	4.2e-26	0.20	-1202.08		PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	TRANSFERASE CALCIUM++, PHOSPHOLIPID BINDING PROTEIN, CALCIUM-BINDING 2 PROTEIN, PHOSPHATIDYL-SERINE, PROTEIN KINASE C
1148	1dhw		240	331	1.5e-22	0.20	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE CALB DOMAIN; HYDROLASE, C2 DOMAIN, CALB

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1148	1rlw		90	181	8.4e-16	0.21	-1202.08		PHOSPHOLIPASE A2; CHAIN: NULL;	DOMAIN HYDROLASE CALB DOMAIN; HYDROLASE, C2 DOMAIN, CALB DOMAIN
1148	3rpb	A	227	340	6e-24	0.05	-1202.08		RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
1148	3rpb	A	72	191	4.2e-20	0.10	-1202.08		RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
1149	1ccx		8	53	0.0006	1.11	-1202.08		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1149	1kap	P	8	53	0.0015	1.00	-1202.08		ALKALINE PROTEASE; IKAP 4 CHAIN: P; IKAP 5 TETRAPEPTIDE (GLY SER ASN SER); IKAP 9 CHAIN: I; IKAP 10	ZINC METALLOPROTEASE P. AERUGINOSA ALKALINE PROTEASE; IKAP 6 CALCIUM BINDING PROTEIN IKAP 19
1149	1qg4	A	24	53	0.003	2.03	-1202.08		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1149	1tal		3	53	3e-05	1.40	-1202.08		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1149	1tal		3	70	0.003	0.91	-1202.08		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1149	1tal		8	63	0.00045	1.19	-1202.08		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1150	1cex		8	67	9e-06	1.25	-1202.08		CUTINASE; CHAIN: NULL;	PROTEINASE
1150	1ga6	A	8	67	0.0006	0.92	-1202.08		SERINE-CARBOXYL PROTEINASE; CHAIN: A; FRAGMENT OF TYROSTATIN; CHAIN: I;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1150	1qq4	A	24	53	0.003	2.03	-1202.08		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1150	1tal		24	67	3e-05	1.57	-1202.08		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1150	1tal		3	63	3e-07	1.27	-1202.08		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1152	1f88	A	107	436	1.5e-15			61.82	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1152	1f88	B	107	434	4.5e-13			68.15	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1153	1etj	A	24	362	5.6e-52	0.54	-1202.08		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
1153	1etj	A	480	761	1.4e-55	0.10	-1202.08		TRANSCRIPTIONAL REPRESSOR TUP1;	TRANSCRIPTION INHIBITOR BETA-PROPELLER

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1153	1ej	A	725	931	2.8e-21	0.11	-1202.08		CHAIN: A, B, C; TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
1153	1ej	A	72	431	4.2e-50	0.16	-1202.08		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
1153	1got	B	17	363	1.4e-58	0.36	-1202.08		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
1153	1got	B	513	802	1.1e-43	0.10	-1202.08		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
1153	1got	B	552	887	4.2e-34	0.10	-1202.08		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
1153	1got	B	644	936	2.8e-28	0.31	-1202.08		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										SUBUNIT, COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
1158	1b6l	A	176	280	4.2e-57	0.22	-1202.08		RETROPEPSIN; CHAIN: A, B;	HYDROLASE/HYDROLASE INHIBITOR HIV-1 PR. COMPLEX (ACID PROTEINASE/PEPTIDE)
1158	1bai	A	162	279	2.8e-17			55.92	ROUS SARCOMA VIRUS PROTEASE; CHAIN: A, B; INHIBITOR; CHAIN: C;	COMPLEX (PROTEASE/INHIBITOR) HUMAN IMMUNODEFICIENCY VIRUS PROTEASE, ROUS SARCOMA VIRUS 2 PROTEASE, CRYSTAL STRUCTURES, PROTEIN-MEDIATED INTERACTION, 3 VIRAL MATURATION, COMPLEX (PROTEASE/INHIBITOR) HEADER
1158	1bai	A	1	111	1.4e-24			52.01	ROUS SARCOMA VIRUS PROTEASE; CHAIN: A, B; INHIBITOR; CHAIN: C;	COMPLEX (PROTEASE/INHIBITOR) HUMAN IMMUNODEFICIENCY VIRUS PROTEASE, ROUS SARCOMA VIRUS 2 PROTEASE, CRYSTAL STRUCTURES, PROTEIN-MEDIATED INTERACTION, 3 VIRAL MATURATION, COMPLEX (PROTEASE/INHIBITOR) HEADER
1158	1bwb	A	176	280	1.4e-60	0.15	-1202.08		HIV-1 PROTEASE; CHAIN: A, B;	HYDROLASE HIV-1 PROTEASE, HYDROLASE
1158	1c6x	A	176	280	2.8e-60	0.42	-1202.08		PROTEASE; CHAIN: A, B;	HYDROLASE HYDROLASE
1158	1daz	C	176	280	5.6e-58	0.29	-1202.08		PEPTIDE INHIBITOR; CHAIN: A, B; HIV-1 PROTEASE (RETROPEPSIN); CHAIN: C, D;	HYDROLASE HIV-1 PROTEASE, MUTANT, DIMER, INHIBITOR, OCCUPANCY
1158	1dun		44	161	7.5e-23			57.26	DEOXYURIDINE 5'-	HYDROLASE DUTPASE, DUTP

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									TRIPHOSPHATE NUCLEOTIDOHYDROLA SE; CHAIN: NULL;	PYROPHOSPHATASE; HYDROLASE, DUTPASE, EIAY, TRIMERIC ENZYME, ASPARTYL PROTEASE
1158	1dun		58	150	7.5e-23	0.91	-1202.08		DEOXYURIDINE 5'- TRIPHOSPHATE NUCLEOTIDOHYDROLA SE; CHAIN: NULL;	HYDROLASE DUTPASE, DUTP PYROPHOSPHATASE; HYDROLASE, DUTPASE, EIAY, TRIMERIC ENZYME, ASPARTYL PROTEASE
1158	1dun		9	126	1.4e-18			56.78	DEOXYURIDINE 5'- TRIPHOSPHATE NUCLEOTIDOHYDROLA SE; CHAIN: NULL;	HYDROLASE DUTPASE, DUTP PYROPHOSPHATASE; HYDROLASE, DUTPASE, EIAY, TRIMERIC ENZYME, ASPARTYL PROTEASE
1158	1euw	A	1	125	7e-22			61.05	DEOXYURIDINE 5'- TRIPHOSPHATE NUCLEOTIDOHYDROLA SE; CHAIN: A;	HYDROLASE DUTPASE; JELLY ROLL, MERCURY DERIVATIVE
1158	1euw	A	32	160	4.5e-18			62.42	DEOXYURIDINE 5'- TRIPHOSPHATE NUCLEOTIDOHYDROLA SE; CHAIN: A;	HYDROLASE DUTPASE; JELLY ROLL, MERCURY DERIVATIVE
1158	1euw	A	58	150	4.5e-18	0.76	-1202.08		DEOXYURIDINE 5'- TRIPHOSPHATE NUCLEOTIDOHYDROLA SE; CHAIN: A;	HYDROLASE DUTPASE; JELLY ROLL, MERCURY DERIVATIVE
1158	1f7d	A	43	157	1.2e-22			62.82	POL POLYPROTEIN; CHAIN: A, B;	VIRUS/VIRAL PROTEIN EIGHT STRANDED BETA-BARREL
1158	1f7d	A	58	150	1.2e-22	0.83	-1202.08		POL POLYPROTEIN; CHAIN: A, B;	VIRUS/VIRAL PROTEIN EIGHT STRANDED BETA-BARREL
1158	1f7d	A	8	122	1.4e-21			62.45	POL POLYPROTEIN; CHAIN: A, B;	VIRUS/VIRAL PROTEIN EIGHT STRANDED BETA-BARREL
1158	1f7r	A	43	180	1.5e-24			73.37	POL POLYPROTEIN; CHAIN: A;	VIRUS/VIRAL PROTEIN EIGHT STRANDED BETA BARREL

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1158	1f7r	A	58	166	1.5e-24	0.29	-1202.08		POL POLYPROTEIN; CHAIN: A;	PROTEIN
1158	1f7r	A	8	136	2.8e-26			71.47	POL POLYPROTEIN; CHAIN: A;	VIRUS/VIRAL PROTEIN EIGHT STRANDED BETA BARREL PROTEIN
1158	1fmb		176	280	2.8e-13			51.30	EIAV PROTEASE; CHAIN: NULL;	HYDROLASE (ACID PROTEINASE) HYDROLASE
1158	1g6l	A	170	280	2.8e-63	-0.00	-1202.08		HIV-1 PROTEASE; CHAIN: A;	HYDROLASE HYDROLASE
1158	1g6l	A	93	280	2.8e-63			53.13	HIV-1 PROTEASE; CHAIN: A;	HYDROLASE HYDROLASE
1158	1hvc		148	280	1.4e-63	0.13	-1202.08		HYDROLASE (ACID PROTEASE) HIV-1 PROTEASE (TETHERED DIMER LINKED BY 1HVC 3 GLY-GLY-SER-SER-GLY) COMPLEXED WITH A-76928 1HVC 4	
1158	1hvc		60	280	1.4e-63			54.32	HYDROLASE (ACID PROTEASE) HIV-1 PROTEASE (TETHERED DIMER LINKED BY 1HVC 3 GLY-GLY-SER-SER-GLY) COMPLEXED WITH A-76928 1HVC 4	
1158	1ida	A	176	279	2.8e-37	0.08	-1202.08		HYDROLASE (ACID PROTEINASE) HUMAN IMMUNODEFICIENCY	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									VIRUS TYPE 2 (HIV-2) PROTEASE IIDA 3 COMPLEXED WITH THE INHIBITOR BILA 1906 CONTAINING THE IIDA 4	
									HYDROXYETHYLAMINE DIPEPTIDE ISOSTERE IIDA 5	
1158	1sip		176	280	1.4e-39	0.29	-1202.08		HYDROLASE(ACID PROTEINASE) SIMIAN IMMUNODEFICIENCY VIRUS (SIV) PROTEINASE 1SIP 3 (SIV MAC251-32H ISOLATE) (E.C.3.4.23.-) 1SIP 4	
1167	1a8q		18	281	8.5e-47	-0.33	0.06		BROMOPEROXIDASE A1; CHAIN: NULL;	HALOPEROXIDASE CHLOROPEROXIDASE A1, HALOPEROXIDASE A1; HALOPEROXIDASE, OXIDOREDUCTASE
1167	1a8s		16	280	1.4e-46	-0.21	0.11		CHLOROPEROXIDASE F; CHAIN: NULL;	HALOPEROXIDASE HALOPEROXIDASE F; HALOPEROXIDASE, OXIDOREDUCTASE, PROPPONATE COMPLEX
1167	1c4x	A	13	270	3.4e-37	-0.33	0.03		2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE CHAIN: A;	HYDROLASE BPHD, HYDROLASE, PCB DEGRADATION
1167	1c4x	A	6	282	3.4e-37			51.46	2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE CHAIN: A;	HYDROLASE BPHD, HYDROLASE, PCB DEGRADATION
1167	1cqw	A	13	281	1.2e-35	-0.34	0.01		HALOALKANE DEHALOGENASE; 1-CHLOROHEXANE	HYDROLASE A/B HYDROLASE FOLD, DEHALOGENASE I-S BOND

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1214	1ev2	E	26	110	3.40E-17	-0.29	0.09		CHAIN: A; FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1214	1ev2	G	26	110	3.40E-17	-0.61	0.05		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1214	1f2q	A	18	107	3.40E-18	0.04	0.93		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1214	1f6a	A	18	106	5.10E-17	0.31	0.9		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IGE EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1214	1fcg	A	9	103	1.00E-17	-0.14	0.23		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32, FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1214	1fhl	A	6	103	3.40E-16	-0.17	0.93		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1214	2fcb	A	9	105	8.50E-19	-0.14	0.17		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
1247	1ee4	A	130	442	1.4e-09	-0.03	0.88		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	SYSTEM TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
1247	1ee4	A	168	498	9.8e-12	0.33	1.00		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
1247	1ial	A	185	485	2.8e-11	0.11	0.95		IMPORTIN ALPHA; CHAIN: A;	NUCLEAR IMPORT RECEPTOR KARYOPHERIN ALPHA; NUCLEAR IMPORT RECEPTOR, NUCLEAR LOCALIZATION SIGNAL, 2 ARMADILLO REPEATS, AUTOINHIBITION, INTRASTERIC REGULATION
1247	2bet		390	501	0.00011	0.23	0.25		BETA-CATENIN; CHAIN: NULL;	STRUCTURAL PROTEIN ARMADILLO REPEAT, BETA-CATENIN, STRUCTURAL PROTEIN
1247	3bet		224	505	8.4e-11	0.28	0.99		BETA-CATENIN; CHAIN: NULL;	ARMADILLO REPEAT ARMADILLO REPEAT, BETA-CATENIN, CYTOSKELETON
1273	1fhs		2	278	1.7e-62			329.79	SULFUR-SUBSTITUTED RHODANESE; CHAIN: NULL;	TRANSFERASE TRANSFERASE, RHODANESE, SULFURTRANSFERASE
1294	1dx5	I	104	218	5.10E-10	0.02	-0.19		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; BGR-CMK SERINE PROTEINASE, BGF-LIKE DOMAINS, ANTICOAGULANT



793

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H; ANTISTASIN; CHAIN: NULL;	COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
1294	1skz		49	163	5.10E-08	0.01	-0.11			
1294	4mt2		71	127	1.70E-09	0.25	-0.12		METALLOTHIONEIN METALLOTHIONEIN ISOFORM II 4MT2.3	
1301	1bx2	A	133	183	2.9E-07	-0.58	0.39		HLA-DR2; CHAIN: A, D; HLA-DR2; CHAIN: B, E; HLA-DR2; CHAIN: C, F;	IMMUNE SYSTEM HLA-DR2, MYELIN BASIC PROTEIN, MULTIPLE SCLEROSIS, 2 AUTOIMMUNITY, IMMUNE SYSTEM
1301	1cqlk	A	132	182	4.3E-07	-0.56	0.05		CH3 DOMAIN OF MAK33 ANTIBODY; CHAIN: A, B;	IMMUNE SYSTEM CONSTANT DOMAIN, C1-SUBSET, IMMUNOGLOBULIN, IMMUNE SYSTEM
1301	1cvs	C	79	133	0.00048	-0.01	0.04		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1301	1dgi	R	20	190	9.6E-16	-0.60	0.00		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS- RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1301	1duz	A	126	183	4.8e-07	-0.51	0.84		HLA-A*0201; CHAIN: A, D; BETA-2 MICROGLOBULIN; CHAIN: B, E; HTLV-1 OCTAMERIC TAX PEPTIDE; CHAIN: C, F;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD
1301	1e4k	A	135	182	9.6e-08	-0.62	0.17		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C, FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
1301	1f5w	A	32	131	9.6e-09	0.47	0.94		COXSACKIE VIRUS AND ADENOVIRUS RECEPTOR; CHAIN: A, B;	VIRUS/VIRAL PROTEIN RECEPTOR IMMUNOGLOBULIN V DOMAIN FOLD, SYMMETRIC DIMER
1301	1f5w	A	92	132	0.00032	-0.20	0.37		COXSACKIE VIRUS AND ADENOVIRUS RECEPTOR; CHAIN: A, B;	VIRUS/VIRAL PROTEIN RECEPTOR IMMUNOGLOBULIN V DOMAIN FOLD, SYMMETRIC DIMER
1301	1fe2	D	135	182	1.4e-07	-0.48	0.40		IMMUNOGLOBULIN IMMUNOGLOBULIN FC AND FRAGMENT B OF PROTEIN A COMPLEX 1FC2 4	
1301	1iak	A	126	183	9.6e-08	-0.85	0.01		MHC CLASS II-I-AK; CHAIN: A, B, P; HEN EGGWHITE LYSOZYME PEPTIDE	HISTOCOMPATIBILITY ANTIGEN I-AK HISTOCOMPATIBILITY ANTIGEN, MHC, PEPTIDE COMPLEX
1301	1neu		34	133	4.8e-06	0.38	0.03		MYELIN P0 PROTEIN; CHAIN: NULL;	STRUCTURAL PROTEIN MYELIN, STRUCTURAL PROTEIN, GLYCOPROTEIN, TRANSMEMBRANE, PHOSPHORYLATION, IMMUNOGLOBULIN FOLD,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										SIGNAL, MYELIN 2 MEMBRANE ADHESION MOLECULE
1333	1alh	A	142	221	3.4e-27	-0.80	0.12		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1333	1alh	A	169	249	1.2e-29	0.01	0.88		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1333	1b6g		385	480	0.0016	0.76	0.74		HALOALKANE DEHALOGENASE; CHAIN: NULL;	HYDROLASE HYDROLASE, HALOALKANE DEHALOGENASE, ALPHA/BETA-HYDROLASE
1333	1mey	C	141	221	1.7e-45	-0.46	0.03		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1333	1mey	C	168	249	1.5e-48	-0.17	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1333	1mey	C	196	277	5.1e-50	0.21	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1333	1mey	C	224	305	8.5e-51	0.19	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FINGER PROTEIN; CHAIN: C, F, G;	INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1333	1mey	C	252	333	3.4e-51	0.40	1.00		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1333	1mey	C	252	334	3.4e-51			107.37	DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1333	1mey	C	308	389	5.1e-51	0.30	1.00		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1333	1mey	C	336	415	8.5e-36	0.10	1.00		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
									PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
1333	1qfm	A	383	576	0.0002	0.23	0.28		TFIIIA; CHAIN: A, D, 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION,
1333	1tft	A	169	314	8.5e-38	-0.13	0.86			

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1333	1tf6	A	171	333	4e-69	0.03	0.92		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	ZINC FINGER PROTEIN COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1333	1tf6	A	196	366	6e-79			110.10	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1333	1tf6	A	197	342	5.1e-38	0.25	0.94		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1333	1tf6	A	197	361	1.6e-78	-0.03	0.98		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1333	1tf6	A	225	371	6.8e-39	0.23	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1333	1tt6	A	225	389	6e-79	0.09	0.89		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	ZINC FINGER PROTEIN
1333	1tt6	A	253	389	1.7e-36	0.22	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2
										TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1333	1tt6	A	281	395	8e-49	0.28	0.86		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2
										TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1333	1tt6	A	281	471	1.4e-32	-0.28	0.09		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2
										TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1333	1ubd	C	143	249	3.4e-33	-0.58	0.54		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1333	Iubd	C	171	277	8e-50	-0.10	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	(TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1333	Iubd	C	176	277	5.1e-35	-0.02	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1333	Iubd	C	194	306	2e-60			90.99	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1333	Iubd	C	222	333	2e-60	0.18	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1333	Iubd	C	232	333	3.4e-35	0.03	0.98		YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1333	1ubd	C	278	389	8e-60	0.18	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1333	1ubd	C	288	389	5.1e-35	0.28	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1333	2gli	A	110	248	1.7e-34	-0.15	0.27		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1333	2gli	A	169	307	6e-64	0.12	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1333	2gli	A	176	304	1e-34	0.20	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1333	2gli	A	196	335	4e-77			98.74	ZINC FINGER PROTEIN GLI; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GL, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1333	2gli	A	196	363	4e-77	0.14	0.81		ZINC FINGER PROTEIN GLI; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GL, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1333	2gli	A	252	390	1e-76	0.36	0.94		ZINC FINGER PROTEIN GLI; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GL, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1333	2gli	A	260	388	3.4e-34	0.08	0.94		ZINC FINGER PROTEIN GLI; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GL, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1334	1a0f	A	154	371	1.7e-24			55.32	GLUTATHIONE S-TRANSFERASE; CHAIN: A, B;	TRANSFERASE GST, GLUTATHIONE TRANSFERASE; TRANSFERASE, GLUTATHIONE CONJUGATION, DETOXIFICATION,
1334	1aw9		160	374	1.2e-42	0.03	0.72		GLUTATHIONE S-TRANSFERASE III; CHAIN: NULL;	TRANSFERASE TRANSFERASE, HERBICIDE DETOXIFICATION
1334	1axd	A	160	366	5.1e-36	0.05	0.21		GLUTATHIONE S-TRANSFERASE I; CHAIN: A, B;	COMPLEX (TRANSFERASE/LIGAND) COMPLEX (TRANSFERASE/LIGAND),
1334	1aem	A	151	378	1e-36	0.22	0.23		LACTOYLGLUTATHION E; CHAIN: C, D	TRANSFERASE, HERBICIDE 2 DETOXIFICATION HEADER
1334	1f3a	A	159	385	3.4e-38	0.09	0.75		GLUTATHIONE-S-TRANSFERASE; CHAIN: A;	TRANSFERASE GST, GLUTATHIONE CONJUGATING, PUTATIVE OXIDOREDUCTASE
1334	1f3a	A	159	385	3.4e-38	0.09	0.75		GLUTATHIONE S-TRANSFERASE	TRANSFERASE GLUTATHIONE S-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									TRANSFERASE YA CHAIN: A, B;	TRANSFERASE, GLUTATHIONE
1334	1gsd	A	160	375	3.4e-35	0.09	0.86		GLUTATHIONE TRANSFERASE A1-1; IGSD 5 CHAIN: A, B; IGSD 6	TRANSFERASE (GLUTATHIONE)
1334	1gse	A	160	385	1.7e-36	0.11	0.75		GLUTATHIONE TRANSFERASE; 1GSE 6 CHAIN: A, B; 1GSE 7	TRANSFERASE (GLUTATHIONE) A1-1 1GSE 19
1334	1hna		160	380	3.4e-37	0.20	0.80		TRANSFERASE (GLUTATHIONE) GLUTATHIONE S-TRANSFERASE (HUMAN, CLASS MU) (GSTM2-2) 1HNA 3 FORM A (E.C.2.5.1.18) MUTANT WITH TRP 214 REPLACED BY PHE 1HNA 4 (W214F) 1HNA 5	
1334	6gsv	A	160	381	1.7e-37	0.04	0.33		MU CLASS GLUTATHIONE S-TRANSFERASE OF ISOENZYME CHAIN: A, B;	GLUTATHIONE TRANSFERASE RAT GST; GLUTATHIONE TRANSFERASE, ISOENZYME 3-3, T13S MUTANT
1348	1ad0	B	19	231	3.40E-70			88.97	FAB FRAGMENT, ANTIBODY A5B7; CHAIN: A, B, C, D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB FRAGMENT
1348	1afv	H	258	443	1.70E-50	0.13	0		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN) HIV-1 CA, HIV CA, HIV P24, P24; FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN), HIV, CAPSID PROTEIN, 2 P24
1348	1aif	H	15	230	5.10E-71	0.1	0.01		ANTI-IDiotYPIC FAB	IMMUNOGLOBULIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									409.5.3 (GG2A) FAB; CHAIN: A, B, L, H	IMMUNOGLOBULIN, C REGION, V REGION
1348	1baf	L	146	328	6.80E-16	-0.04	0.17		IMMUNOGLOBULIN FAB FRAGMENT OF MURINE MONOCLONAL ANTIBODY AN02 COMPLEX IBAF 3 WITH ITS HAPTEN (2,2,6,6-TETRAMETHYL-1-PIPERIDINYLOXY-1BAF 4 DINITROPHENYL) IBAF 5	
1348	1bfo	B	19	230	1.50E-67			80.18	CAMPATH-1G ANTIBODY; CHAIN: A, B, C, D, E, F, G, H;	ANTIBODY ANTIBODY, FAB, CAMPATH-1G, CD52
1348	1bli	A	143	528	6.80E-26	0.01	0.7		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1348	1bli	A	44	443	1.40E-46			118.73	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1348	1bii	A	61	443	1.40E-46	0.23	0.57		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1348	1bin	B	15	232	1.70E-72	0.06	0.25		MONOCLONAL ANTIBODY MRK-16 (LIGHT CHAIN); CHAIN: A, C; MONOCLONAL ANTIBODY MRK-16 (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM IMMUNOGLOBULIN, IMMUNE SYSTEM
1348	1bin	B	19	231	1.70E-72			79.05	MONOCLONAL ANTIBODY MRK-16 (LIGHT CHAIN); CHAIN:	IMMUNE SYSTEM IMMUNOGLOBULIN, IMMUNE SYSTEM

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A, C; MONOCLONAL ANTIBODY MRK-16 (HEAVY CHAIN); CHAIN: B, D;	
1348	1cl7	I	158	235	3.40E-19	0.35	0.64		IGG1 ANTIBODY 1696 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 1696 (VARIABLE HEAVY CHAIN); CHAIN: H; IGG1 ANTIBODY 1696 (CONSTANT HEAVY CHAIN); CHAIN: I;	IMMUNE SYSTEM IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1 FAB FRAGMENT, CROSS-REACTIVITY, HIV1 PROTEASE, ENZYME 2 INHIBITION, IMMUNOGLOBULIN
1348	1clo	H	15	234	1.70E-72	-0.01	0.22		A5B7 MONOCLONAL ANTIBODY; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB- FRAGMENT
1348	1clo	H	19	233	1.70E-72			89.82	A5B7 MONOCLONAL ANTIBODY; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB- FRAGMENT
1348	1clz	H	19	233	3.40E-69			83.21	IGG FAB (IGG3, KAPPA); CHAIN: L, H;	IMMUNOGLOBULIN MBR96 FAB (IMMUNOGLOBULIN); IMMUNOGLOBULIN C REGION, GLYCOPROTEIN, TRANSMEMBRANE
1348	1cs6	A	140	540	8.50E-38	0.02	0.49		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1348	1cs6	A	255	576	1.00E-30	-0.01	0.48		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1348	1cs6	A	40	444	5.10E-63	0.36	0.86		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1348	1cvs	C	137	341	3.40E-20	0.03	-0.01		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										FACTOR/GROWTH FACTOR RECEPTOR
1348	1cvs	D	137	341	1.40E-21	0.19	0.19		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1348	1d5i	H	258	454	6.80E-47	0.09	0.21		CHIMERIC GERMLINE PRECURSOR OF OXY- COPE CHAIN: L; CHIMERIC GERMLINE PRECURSOR OF OXY- COPE CHAIN: H;	IMMUNE SYSTEM IMMUNE SYSTEM
1348	1dee	A	143	327	5.10E-18	-0.09	0.01		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
1348	1dfb	H	258	454	1.70E-47	-0.08	0.83		IMMUNOGLOBULIN 3D6 FAB 1DFB 3	
1348	1dfb	L	143	327	1.00E-19	-0.04	0.16		IMMUNOGLOBULIN 3D6 FAB 1DFB 3	
1348	1dgi	R	23	340	8.20E-38	0.13	0.01		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1348	1dn2	A	247	437	1.20E-39	0.07	-0.18		IMMUNOGLOBULIN LAMBDA HEAVY	IMMUNE SYSTEM FC IGG PHAGE DISPLAY PEPTIDE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: CHAIN: A, B; ENGINEERED PEPTIDE; CHAIN: E, F;	
1348	1dzb	A	15	215	8.50E-58	0.23	-0.13		SCFV FRAGMENT 1F9; CHAIN: A, B; TURKEY EGG-WHITE LYSOZYME C; CHAIN: X, Y;	COMPLEX (ANTIBODY ANTIGEN) 1,4-BETA-N-ACETYLMURAMIDASE C; SINGLE-DOMAIN ANTIBODY, TURKEY EGG-WHITE LYSOZYME, 2 ANTIBODY-PROTEIN COMPLEX, SINGLE-CHAIN FV FRAGMENT
1348	1dzb	A	256	425	3.40E-35	0.09	0.22		SCFV FRAGMENT 1F9; CHAIN: A, B; TURKEY EGG-WHITE LYSOZYME C; CHAIN: X, Y;	COMPLEX (ANTIBODY ANTIGEN) 1,4-BETA-N-ACETYLMURAMIDASE C; SINGLE-DOMAIN ANTIBODY, TURKEY EGG-WHITE LYSOZYME, 2 ANTIBODY-PROTEIN COMPLEX, SINGLE-CHAIN FV FRAGMENT
1348	1evt	C	137	341	1.70E-21	0.22	0.04		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFI; FGFR1; IMMUNOGLOBULIN (G) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1348	1f58	H	258	444	1.70E-49	-0.17	0.11		IGG1 ANTIBODY 58.2 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 58.2 (HEAVY CHAIN); CHAIN: H; EXTERIOR MEMBRANE GLYCOPROTEIN(GP120); CHAIN: P;	IMMUNE SYSTEM FAB 58.2; FAB 58.2; V3 LOOP; IMMUNOGLOBULIN, FAB, HIV-1, GP120, V3, IMMUNE SYSTEM
1348	1fai	H	258	440	1.50E-50	0.19	0.34		IMMUNOGLOBULIN FAB FRAGMENT FROM A MONOCLONAL ANTI-ARSONATE ANTIBODY,	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									R19.9 1FAI 3 (IGG2B,KAPPA) 1FAI 4 COMPLEX (ANTIBODY/ANTIGEN) FAB FRAGMENT OF THE MONOCLONAL ANTIBODY F9.13.7 (GG1) 1FBI 3 COMPLEXED WITH LYSOZYME (E.C.3.2.1.17) 1FBI 4	
1348	1fbi	H	258	440	1.70E-51	-0.06	0.12			
1348	1fcg	A	237	443	4.10E-20	0.26	0.51		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32: FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1348	1feg	H	19	233	5.10E-67			79.18	IMMUNOGLOBULIN/VIR US HEMAGGLUTININ IGG2A FAB FRAGMENT (FAB 26/9) COMPLEXED WITH INFLUENZA 1FRG 3 HEMAGGLUTININ HAI (STRAIN X47) (RESIDUES 101 - 108) 1FRG 4	
1348	1fvd	A	143	327	1.40E-17	-0.13	0.24		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
1348	1fvd	B	258	454	6.80E-48	0.02	0.59		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
1348	1hil	B	19	233	6.80E-71			79.17	IMMUNOGLOBULIN IGG2A FAB FRAGMENT (FAB 17/9) 1HIL 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1348	1iai	H	13	231	1.70E-68			80.32	IDIOTYPIC FAB 730.1.4 (IGG1) OF VIRUS 11AI 5 CHAIN: L, H, 11AI 7 ANTI-IDIOTYPIC FAB 409.5.3 (IGG2A); 11AI 9 CHAIN: M, I 11AI 10	COMPLEX (IMMUNOGLOBULIN IGG1/IGG2A)
1348	1igc	H	16	235	1.40E-72	-0.02	0.07		COMPLEX (ANTIBODY/BINDING PROTEIN) IGG1 FAB FRAGMENT COMPLEXED WITH PROTEIN G (DOMAIN III) IIGC 5 PROTEIN G, STREPTOCOCCUS IIGC 15	
1348	1igc	H	258	444	3.40E-51	0.11	0.03		COMPLEX (ANTIBODY/BINDING PROTEIN) IGG1 FAB FRAGMENT COMPLEXED WITH PROTEIN G (DOMAIN III) IIGC 5 PROTEIN G, STREPTOCOCCUS IIGC 15	
1348	1igt	B	15	439	0	-0.02	0.13		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1348	1igt	B	23	441	0			124.15	IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1348	1igy	B	14	441	0			114.03	IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
1348	1ikf	H	256	444	3.40E-51	0.04	-0.08		COMPLEX (ANTIBODY/IMMUNOSUP	



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(PRESSANT) IMMUNOGLOBULIN IGG1-KAPPA ANTIBODY FRAGMENT FAB COMPLEXED IKF 3 WITH CYCLOSPORIN IKF 4	
1348	1tib	B	143	443	2.90E-21	0.12	0.47		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1348	1kel	H	19	230	1.20E-67			78.42	28B4 FAB; CHAIN: L, H;	CATALYTIC ANTIBODY SULFIDE OXIDATION, MONOOXYGENASE, OXYGENATION, FAB, 2 IMMUNOGLOBULIN, CATALYTIC ANTIBODY
1348	1hi1	A	148	336	1.00E-21	-0.08	0.19		LAMBDA III BENCE JONES PROTEIN CLE; CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENCE JONES PROTEIN
1348	1lmk	A	16	215	1.50E-52	0.06	-0.13		IMMUNOGLOBULIN ANTI- PHOSPHATIDYLINOSITOL SPECIFIC PHOSPHOLIPASE C DIABODY 1LMK 3 SYNONYMS: L5MK16 DIABODY, SINGLE-CHAIN FV DIMER 1LMK 4	
1348	1lmk	A	258	428	5.10E-33	0.1	0.34		IMMUNOGLOBULIN ANTI- PHOSPHATIDYLINOSITO	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									L.SPECIFIC PHOSPHOLIPASE C DIABODY 1LMK 3 SYNONYMS: L5MK16 DIABODY, SINGLE-CHAIN FV DIMER 1LMK 4	
1348	Imco	H	147	527	5.10E-35	-0.24	0.09		IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION IMCO 3	
1348	Imco	H	18	437	8.50E-98	0.03	0.27		IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION IMCO 3	
1348	Imco	H	20	442	8.50E-98			101.92	IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION IMCO 3	
1348	Imcp	H	256	429	5.10E-35	0.29	0.51		IMMUNOGLOBULIN FAB FRAGMENT (MC/PCS603) IMCP 4	
1348	Imfa		148	352	6.80E-24	0.24	0.09		IMMUNOGLOBULIN FV FRAGMENT (MURINE SE155-4) COMPLEX WITH THE TRISACCHARIDE: IMFA 3 ALPHA-D-GALACTOSE(1-2)[ALPHA-D-ABEQUOSE(1-3)]ALPHA-	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									IMFA 4 D-MANNOSE (P1-OME) (PART OF THE CELL-SURFACE CARBOHYDRATE IMFA 5 OF PATHOGENIC SALMONELLA) IMFA 6	
1348	Inca	H	15	233	5.10E-67			78.54	HYDROLASE(O-GLYCOSYL) N9 NEURAMINIDASE-NC41 (E.C.3.2.1.18) COMPLEX WITH FAB INCA 3	
1348	Infd	F	19	233	1.70E-70			83.51	N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN)
1348	Ingp	H	15	233	1.70E-73	0.1	0.12		NIG9 (GGI=LAMBDA=); CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN,
1348	Inqb	A	258	425	1.70E-33	0.02	0.69		SINGLE-CHAIN ANTIBODY FRAGMENT; CHAIN: A, C;	IMMUNOGLOBULIN VARIABLE HEAVY (VH) DOMAIN, VARIABLE LIGHT (VL) ANTIBODY FRAGMENT, MULTIVALENT ANTIBODY, DOMAIN 2 SWAPPING, IMMUNOGLOBULIN
1348	lpsk	H	258	438	6.80E-41	0.05	0.12		ANTIBODY; CHAIN: L, H;	IMMUNOGLOBULIN FAB, GD2-GANGLIOSIDE, CARBOHYDRATE, MELANOMA, IMMUNOGLOBULIN
1348	lqkz	H	16	233	6.80E-72	-0.09	0.12		ANTIBODY; CHAIN: H, L; PROTEIN G-PRIME; CHAIN: A; MAJOR OUTER MEMBRANE PROTEIN P1.16; CHAIN: P;	IMMUNE SYSTEM FAB, PORIN, NEISSERIA MENINGITIDIS, PORIN
1348	lqok	A	15	220	3.40E-56	0.04	-0.13		MFE-23 RECOMBINANT	IMMUNOGLOBULIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ANTIBODY FRAGMENT; CHAIN: A;	IMMUNOGLOBULIN, SINGLE-CHAIN FV, ANTI-CARCINOEMBRYONIC 2 ANTIGEN
1348	1gok	A	256	430	1.70E-34	0.36	0.41		MFE-23 RECOMBINANT ANTIBODY FRAGMENT; CHAIN: A;	IMMUNOGLOBULIN IMMUNOGLOBULIN, SINGLE-CHAIN FV, ANTI-CARCINOEMBRYONIC 2 ANTIGEN
1348	1t24	B	258	434	5.10E-45	0.24	0.3		IGG3-KAPPA ANTIBODY (LIGHT CHAIN); CHAIN: A, C; IGG3-KAPPA ANTIBODY (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM PRELIMINARY, IMMUNE SYSTEM
1348	1sbs	H	15	232	6.80E-72	0.23	0.54		MONOCLONAL ANTIBODY 3A2; CHAIN: H, L;	MONOCLONAL ANTIBODY MONOCLONAL ANTIBODY, FAB-FRAGMENT, REPRODUCTION
1348	1sbs	H	23	230	6.80E-72			80.32	MONOCLONAL ANTIBODY 3A2; CHAIN: H, L;	MONOCLONAL ANTIBODY MONOCLONAL ANTIBODY, FAB-FRAGMENT, REPRODUCTION
1348	1sbs	H	258	440	1.20E-50	-0.02	0.1		MONOCLONAL ANTIBODY 3A2; CHAIN: H, L;	MONOCLONAL ANTIBODY MONOCLONAL ANTIBODY, FAB-FRAGMENT, REPRODUCTION
1348	1sm3	H	15	233	8.50E-73	-0.11	0		SM3 ANTIBODY; CHAIN: L, H; PEPTIDE EPTOPE; CHAIN: P;	COMPLEX (ANTIBODY/PEPTIDE EPTOPE) ANTIBODY, PEPTIDE ANTIGEN, ANTITUMOR ANTIBODY, 2 COMPLEX (ANTIBODY/PEPTIDE EPTOPE)
1348	1sm3	H	23	233	8.50E-73			79.3	SM3 ANTIBODY; CHAIN: L, H; PEPTIDE EPTOPE; CHAIN: P;	COMPLEX (ANTIBODY/PEPTIDE EPTOPE) ANTIBODY, PEPTIDE ANTIGEN, ANTITUMOR ANTIBODY, 2 COMPLEX (ANTIBODY/PEPTIDE EPTOPE)
1348	1tet	H	20	233	1.00E-65			78.05	IMMUNOGLOBULIN IGG1 MONOCLONAL FAB FRAGMENT (TE33)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1348	1vhp		15	141	1.20E-47	0.11	-0.19		COMPLEX WITH CHOLERA 1TET 3 TOXIN PEPTIDE 3 (CTP3) 1TET 4 VH-P8; CHAIN: NULL;	IMMUNOGLOBULIN NMR, VH DOMAIN, ANTIBODY, HUMAN, IMMUNOGLOBULIN
1348	1wio	A	98	428	8.20E-14	0.02	0.13		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
1348	25c8	L	146	328	1.00E-15	0.01	-0.01		IGG 5C8; CHAIN: L, H;	CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
1348	2fb4	H	258	453	6.80E-47	-0.01	0.33		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB 2FB4 4	
1348	2fbj	H	256	429	1.70E-34	0.19	0.62		IMMUNOGLOBULIN IG*A FAB FRAGMENT (J539) (GALACTAN-BINDING) 2FBJ 3	
1348	2gfb	B	16	233	1.20E-73	0.09	0.53		IMMUNOGLOBULIN IGG2A FAB FRAGMENT (CNJ206) 2GFB 3	
1348	2hlp	H	256	440	6.80E-53	-0.14	0.07		2H1; CHAIN: L, H; PA1; CHAIN: P;	COMPLEX (ANTIBODY/PEPTIDE) ANTIBODY STRUCTURE, CRYPTOCOCUS, PEPTIDE, PHAGE LIBRARY, 2 POLYSACCHARIDE, COMPLEX (ANTIBODY/PEPTIDE)
1348	6fab	H	258	444	6.80E-50	-0.07	0.06		IMMUNOGLOBULIN ANTIGEN-BINDING FRAGMENT OF THE MURINE ANTI-	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PHENYLARSONATE 6FAB 3 ANTIBODY 36-71, FAB 36-71 6FAB 4	
1348	8fab	A	148	336	3.40E-21	0.17	0.29		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN	
									IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
1356	12e8	L	22	222	1.7e-17			74.09	2E8 (IGG1=KAPPA=) ANTIBODY; CHAIN: L, H, M, P;	IMMUNOGLOBULIN IMMUNOGLOBULIN
1356	1a4fj	L	22	222	1.5e-19			73.51	IMMUNOGLOBULIN, DIELS ALDER CATALYTIC ANTIBODY; CHAIN: L, H, A, B;	IMMUNOGLOBULIN ANTIBODY, CATALYTIC ANTIBODY, DIELS ALDER, 2 GERMLINE
1356	1ad0	A	29	222	3.4e-22	0.15	0.35		FAB FRAGMENT, ANTIBODY A5B7; CHAIN: A, B, C, D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB FRAGMENT
1356	1adq	L	25	224	1.4e-26			71.74	IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
1356	1adq	L	28	219	1.4e-26	0.09	0.78		IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
1356	1b2w	L	22	222	1.4e-22			74.84	ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN; CHAIN: H;	ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
1356	1b2w	L	29	222	1.4e-22	0.10	0.80		ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
1356	1b4j	L	29	222	1.2e-21	-0.07	0.31		ANTIBODY; CHAIN: L, H;	IMMUNE SYSTEM ANTIBODY ENGINEERING ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODIES, 2 FAB, X-RAY STRUCTURES, GAMMA-INTERFERON
1356	1b6d	A	22	222	1.7e-22			74.61	IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
1356	1b6d	A	29	222	1.7e-22	-0.03	0.69		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
1356	1b6j	L	22	222	1e-19			80.38	IMMUNOGLOBULIN FAB FRAGMENT OF MONOCLONAL ANTIBODY B72.3 1B3J 3 (MURINE/HUMAN CHIMERA) 1B3J 4	
1356	1b1h	A	21	394	1.5e-55			115.28	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1356	1b1h	A	28	395	1.5e-55	0.01	0.07		HEMOLIN; CHAIN: A, B;	HOMOPHILIC ADHESION
1356	1b1l	L	22	222	1.4e-22			72.47	FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1356	1b1l	L	29	222	1.4e-22	0.02	0.39		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12, VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
1356	1b0g	A	22	222	3.4e-19			73.26	ANTIBODY (CB 4-1); CHAIN: A, B; PEPTIDE; CHAIN: C;	COMPLEX (ANTIBODY/PEPTIDE) POLYSPECIFICITY, CROSS REACTIVITY, FAB-FRAGMENT, PEPTIDE, 2 HIV-1, COMPLEX (ANTIBODY/PEPTIDE)
1356	1b27	A	22	218	3.4e-18			74.03	ANTIBODY R24 (LIGHT CHAIN); CHAIN: A; ANTIBODY R24 (HEAVY CHAIN); CHAIN: B;	IMMUNE SYSTEM ANTIBODY (FAB FRAGMENT), IMMUNE SYSTEM
1356	1cdy		135	315	6.3e-23	0.35	0.64		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
1356	1ce1	L	22	222	1.2e-21			72.17	CAMPATH-1H LIGHT CHAIN; CHAIN: L;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CAMPATH-1H:HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	
1356	1cs6	A	26	395	1.7e-58	0.23	0.55		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1356	1cvs	C	122	315	1.2e-42	0.32	0.52		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1cvs	C	240	395	1.7e-35	0.07	-0.18		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1cvs	C	29	120	6.8e-13	-0.07	0.03		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1cvs	C	34	223	1.4e-28	0.03	0.10		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1cvs	D	122	315	1.2e-42	0.28	0.47		FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1cvs	D	240	395	3.4e-37	0.16	-0.14		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1cvs	D	29	120	1e-11	-0.14	0.33		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1cvs	D	34	223	1e-26	-0.12	0.37		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1356	1d5l	L	29	222	3.4e-22	-0.08	0.65		CHIMERIC GERMLINE PRECURSOR OF OXY-COPE CHAIN: L; CHIMERIC GERMLINE PRECURSOR OF OXY-COPE CHAIN: H;	IMMUNE SYSTEM IMMUNE SYSTEM
1356	1dee	A	29	222	1.4e-22	0.10	0.43		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
1356	1dfb	L	22	222	1.2e-21			76.72	IMMUNOGLOBULIN 3D6 FAB 1DFB 3	
1356	1dgi	R	29	315	5.1e-43	-0.29	0.27		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PYR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS- RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1356	1epf	A	118	305	4.2e-28	0.47	0.63		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1356	1epf	A	128	299	5.1e-22	0.34	1.00		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1356	1epf	A	235	379	1.7e-21	0.05	-0.09		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1356	1epf	A	27	225	5.1e-26	0.25	0.11		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1356	1ev2	E	123	315	6.8e-39	-0.11	0.05		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I- SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1ev2	E	240	395	1.5e-33	0.06	-0.17		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1ev2	E	34	223	6.8e-24	0.24	0.15		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1ev2	G	123	319	8.5e-43	0.29	0.23		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
									F, G, H;	DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1ev2	G	240	395	1.7e-36	0.10	-0.07		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
									F, G, H;	DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1ev2	G	29	124	1.7e-11	-0.03	0.24		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
									F, G, H;	DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1ev2	G	34	223	1e-24	0.14	0.42		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
									F, G, H;	DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1ev1	C	122	315	1.7e-41	0.23	0.39		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1evt	C	29	120	1.7e-11	-0.08	0.10		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1evt	C	34	223	5.1e-25	-0.07	0.22		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1356	1f2q	A	119	318	2.1e-19	0.18	0.35		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON R-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1356	1fcg	A	123	315	6.3e-23	0.23	0.55		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1356	1fcg	A	228	395	3.4e-18	0.02	-0.19		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1356	1fhg	A	225	315	8.5e-18	0.25	0.94		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
1356	1fhg	A	31	120	3.4e-14	0.14	0.30		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL

822

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1356	1fvd	A	22	222	3.4e-22			73.26	IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 IFVD 3	BARREL
1356	1fvd	A	29	222	3.4e-22	0.12	0.55		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 IFVD 3	
1356	1ge1	L	22	222	3.4e-21			73.99	ENVELOPE PROTEIN GP120; CHAIN: G; CD4; CHAIN: C; ANTIBODY 17B; CHAIN: L, H;	COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB), HIV-1 EXTERIOR 2 ENVELOPE GP120, T-CELL SURFACE GLYCOPROTEIN CD4, 3 ANTIGEN-BINDING FRAGMENT OF HUMAN IMMUNOGLOBULIN 17B, 4 GLYCOSYLATED PROTEIN
1356	1hng	A	135	317	4.2e-26	0.32	0.03		T LYMPHOCYTE ADHESION GLYCOPROTEIN CD2 (RAT) IHNG 3	
1356	1igt	B	23	393	1e-18			74.95	IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1356	1ib	B	145	393	5.1e-19	-0.07	0.09		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1356	1ib	B	32	319	2.1e-24			81.96	INTERLEUKIN-1 BETA;	COMPLEX

823

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	(IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1356	1iib	B	34	315	2.1e-24	0.04	0.06		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1356	1koa		26	120	1.7e-11	-0.35	0.12		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
1356	1ihl	A	28	219	1.7e-25	0.24	0.98		LAMBDA III BENGE JONES PROTEIN CLE; CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENGE JONES PROTEIN
1356	1mco	H	10	395	8.5e-25			83.75	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION IMCO 3	
1356	1mco	H	25	384	8.5e-25	-0.18	0.33		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (GG1) (MCG) WITH A HINGE DELETION IMCO 3	
1356	1nct		226	315	1.1e-17	0.40	0.16		TTIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1356	1nct		29	121	1.5e-14	0.11	0.05		TTTN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTIN5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
1356	1nfd	E	27	220	5.1e-25	0.38	0.74		N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN)
1356	1osp	L	22	222	1.7e-16			76.74	FAB 184.1; CHAIN: L, H; OUTER SURFACE PROTEIN A; CHAIN: O;	COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN) OSPA; COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN), OUTER SURFACE 2 PROTEIN A COMPLEXED WITH FAB184.1, BORRELIA BURGDORFERI 3 STRAIN B31
1356	1tmm		29	121	1.5e-14	-0.03	0.18		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNNM3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNNM4 1TNNM58	
1356	2feb	A	228	395	3.4e-18	0.01	-0.18		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1356	2feb	A	28	191	8.5e-14	-0.28	0.01		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1356	2fgw	L	29	222	5.1e-23	-0.07	0.29		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION	



825

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
1356	2ncm		27	120	3.4e-12	0.21	0.36		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
1356	3fct	A	22	222	1.7e-21			77.52	METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: A, C; METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: B, D;	IMMUNE SYSTEM METAL CHELATASE, CATALYTIC ANTIBODY, FAB FRAGMENT, IMMUNE 2 SYSTEM
1356	8fab	A	26	221	8.5e-29			73.89	IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
1356	8fab	A	28	219	8.5e-29	0.18	0.94		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
1367	1b1i	A	29	129	4.2e-30	-0.05	0.76		HYDROLASE ANGIOGENIN; CHAIN: A;	HYDROLASE HYDROLASE (VASCULARIZATION)
1367	1b6v	A	21	132	1.7e-43			50.95	RIBONUCLEASE; CHAIN: A, B;	MOLECULAR EVOLUTION MOLECULAR EVOLUTION, RIBONUCLEASE
1367	1b6v	A	27	132	1.7e-43	-0.27	0.98		RIBONUCLEASE; CHAIN:	MOLECULAR EVOLUTION

826

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQEQL D score	Compound	PDB annotation
									A, B;	MOLECULAR EVOLUTION, RIBONUCLEASE
1367	1bstr	A	27	132	5.1e-42	-0.47	0.66		HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE (BOVINE, SEMINAL) (BS-RNASE) 1BSR 3	
1367	1dy5	A	27	132	3.4e-44	0.16	0.93		RIBONUCLEASE A; CHAIN: A, B;	HYDROLASE (PHOSPHORIC DIESTER) RIBONUCLEASE, DEAMIDATION, ULTRA-HIGH RESOLUTION
1367	1qnt	A	24	132	3.4e-41	-0.11	1.00		EOSINOPHIL CATIONIC PROTEIN; CHAIN: A;	RIBONUCLEASE RIBONUCLEASE 3, RNASE 3 EOSINOPHIL, RIBONUCLEASE, CYTOTOXICITY
1367	1rbd		41	132	1.7e-39	-0.51	0.88		HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE S (E.C.3.1.27.5) MUTANT WITH MET 13 REPLACED BY 1RBD 3 ALPHA-AMINO-NORMAL-BUTYRIC ACID (M13ABA) 1RBD 4	
1367	1mf	A	31	130	3.4e-34	0.10	0.96		RIBONUCLEASE 4; CHAIN: A, B;	HYDROLASE RNASE 4; HYDROLASE, RIBONUCLEASE, PHOSPHODIESTERASE
1367	1mu		27	132	8.5e-45	-0.13	0.93		HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE S (E.C.3.1.27.5) (PH.5.5) 1RNU 3	
1367	1tra	A	28	132	1.7e-41	0.00	0.96		RIBONUCLEASE; CHAIN: A;	HYDROLASE RNASE A; HYDROLASE(PHOSPHORIC DIESTER), RIBONUCLEASE

827

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1367	1srn	A	27	114	1.7e-39	-0.20	0.90		HYDROLASE (NUCLEIC ACID,RNA) SEMISYNTHETIC RIBONUCLEASE A (R*NASE 1-118(COLON)111-124) ISRN 3 (E.C.3.1.27.5) ISRNA 1 ISRNA 2	
1388	1awq	A	42	193	1.7e-75			166.81	CYCLOPHILIN A; CHAIN: A; PEPTIDE FROM THE HIV-1 CAPSID PROTEIN; CHAIN: B;	COMPLEX (ISOMERASE/PEPTIDE) COMPLEX (ISOMERASE/PEPTIDE), CYCLOPHILIN A, HIV-1 CAPSID, 2 PSEUDO-SYMMETRY
1388	1eyn	A	39	194	3.4e-65			225.00	CYCLOPHILIN B; ICYN 6 CHAIN: A; ICYN 7 ID-(CHOLINYL)ALA]8-CYCLOSPORIN; ICYN 10 CHAIN: C; ICYN 11	COMPLEX (ISOMERASE/IMMUNOSUPPRESS ANT) CYCLOSPORIN, ISOMERASE, ROTAMASE, SIGNAL ICYN 19
1392	1tal		28	112	0.006	0.84	0.19		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1393	1l88	A	101	168	3.4e-06	-0.92	0.01		RHODOP SIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1423	1b6v	A	32	133	8.5e-49	0.31	1.00		RIBONUCLEASE; CHAIN: A, B;	MOLECULAR EVOLUTION MOLECULAR EVOLUTION, RIBONUCLEASE
1423	1b6v	A	32	148	8.5e-49			77.74	RIBONUCLEASE; CHAIN: A, B;	MOLECULAR EVOLUTION MOLECULAR EVOLUTION, RIBONUCLEASE
1423	1bst	A	32	133	1e-47	0.01	0.87		HYDROLASE(PHOSPHO	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									RIC DIESTER, RNA) RIBONUCLEASE (BOVINE, SEMINAL) (BS- RNASE) IBSR 3	
1423	lbsr	A	32	148	1e-47			66.58	HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE (BOVINE, SEMINAL) (BS- RNASE) IBSR 3	
1423	ldy5	A	32	133	5.1e-47	0.38	1.00		RIBONUCLEASE A; CHAIN: A, B;	HYDROLASE (PHOSPHORIC DIESTER) RIBONUCLEASE, DEAMIDATION, ULTRA-HIGH RESOLUTION
1423	lqnt	A	30	133	3.4e-36	0.29	1.00		EOSINOPHIL CATIONIC PROTEIN; CHAIN: A;	RIBONUCLEASE RIBONUCLEASE 3, RNASE 3 EOSINOPHIL, RIBONUCLEASE, CYTOTOXICITY
1423	lrbd		46	133	8.5e-44	0.51	0.96		HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE S (E.C.3.1.27.5) MUTANT WITH MET 13 REPLACED BY IRBD 3 ALPHA-AMINO- NORMAL-BUTYRIC ACID (M13ABA) IRBD 4	
1423	lrbd		46	148	8.5e-44			77.31	HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE S (E.C.3.1.27.5) MUTANT WITH MET 13 REPLACED BY IRBD 3 ALPHA-AMINO- NORMAL-BUTYRIC ACID (M13ABA) IRBD 4	
1423	lmf	A	32	133	1.7e-35	0.36	0.95		RIBONUCLEASE 4;	HYDROLASE RNASE 4;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A, B;	HYDROLASE, RIBONUCLEASE, PHOSPHODIESTERASE
1423	1mrf	A	33	152	1.7e-35			58.15	RIBONUCLEASE 4; CHAIN: A, B;	HYDROLASE RNASE 4; HYDROLASE, RIBONUCLEASE, PHOSPHODIESTERASE
1423	1rmu		32	133	3.4e-49	0.36	0.99		HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE S (E.C.3.1.27.5) (PH 5.5) IRNU 3	
1423	1rmu		32	148	3.4e-49			79.09	HYDROLASE(PHOSPHO RIC DIESTER, RNA) RIBONUCLEASE S (E.C.3.1.27.5) (PH 5.5) IRNU 3	
1423	1rra	A	33	133	5.1e-45	0.04	0.99		RIBONUCLEASE; CHAIN: A;	HYDROLASE RNASE A; HYDROLASE(PHOSPHORIC DIESTER), RIBONUCLEASE
1423	1sm	A	32	133	3.4e-49	0.25	0.99		HYDROLASE (NUCLEIC ACID, RNA) SEMISYNTHETIC RIBONUCLEASE A (R*NASE 1-118(COLON)111-124) ISRN 3 (E.C.3.1.27.5) ISRNA 1 ISRNA 2	
1423	1sm	A	32	134	3.4e-49			80.10	HYDROLASE (NUCLEIC ACID, RNA) SEMISYNTHETIC RIBONUCLEASE A (R*NASE 1-118(COLON)111-124) ISRN 3 (E.C.3.1.27.5) ISRNA 1 ISRNA 2	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1427	1a4y	A	1	166	4e-20	0.16	-0.01		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1427	1a9n	A	1	129	6e-21	0.41	0.64		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1427	1a9n	A	1	63	0.00017	0.07	0.63		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1427	1a9n	A	33	179	4e-20	0.62	0.47		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1427	1a9n	C	1	135	4e-21	0.40	0.37		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1427	1a9n	C	33	179	8e-20	0.51	0.78		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1427	1b1h	A	210	330	4e-10	0.44	0.11		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1427	1b1v		339	424	0.0015	-0.02	0.07		TITIN; CHAIN: NULL;	CONNECTIN A71, CONNECTIN; TITIN, CONNECTIN, FIBRONECTIN TYPE III

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1427	1cs6	A	166	298	1.7e-10	-0.17	0.22		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1427	1cs6	A	210	342	4e-13	0.27	0.24		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1427	1cvs	C	229	360	1.4e-06	0.15	0.16		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1427	1cvs	D	211	299	1e-16	0.50	0.62		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1427	1cvs	D	229	350	5.1e-06	0.24	0.10		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1427	1d0b	A	1	154	3.4e-22	0.61	0.96		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1427	1dce	A	1	115	4e-14	0.37	0.86		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1427	1dce	A	29	135	8.5e-10	0.45	1.00		SUBUNIT; CHAIN: B, D; RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 Å 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1427	1ds9	A	7	140	1.2e-18	0.22	0.65		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1427	1ev2	G	207	302	2e-17	0.57	0.37		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I- SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1427	1ev2	G	222	342	2e-14	0.32	0.22		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGF2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I- SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1427	1evt	C	222	342	2e-14	0.34	0.35		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGF1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I- SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1427	1fhg	A	205	301	1.8e-18	0.38	0.59		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1427	1fhh	A	154	422	4e-07			72.40	FIBRONECTIN; CHAIN: A;	HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
1427	1fhh	A	241	419	4e-07	0.37	0.33		FIBRONECTIN; CHAIN: A;	HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
1427	1fhh	A	318	416	6.8e-06	-0.42	0.43		FIBRONECTIN; CHAIN: A;	HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
1427	1fhh	A	340	442	0.00014	0.05	0.39		FIBRONECTIN; CHAIN: A;	HEPARIN AND INTEGRIN BINDING HEPARIN AND INTEGRIN BINDING
1427	1fo1	A	24	65	6.8e-05	-0.44	0.51		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1427	1fo1	A	2	62	1.7e-05	0.02	0.23		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1427	1fo1	B	24	65	6.8e-05	-0.49	0.43		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1427	1fo1	B	2	62	1.7e-05	0.15	0.76		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1427	1fqv	A	1	146	1.4e-11	0.09	-0.12		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1427	1fs2	A	1	135	4e-16	0.17	-0.05		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	PROTEIN LIGASE LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1427	1mfn		333	419	2e-07	0.01	0.28		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN
1427	1mfn		341	416	1e-06	-0.03	0.53		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN
1427	1nct		209	298	1e-19	0.43	0.92		TITIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTIN5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
1427	1qr4	A	230	415	6e-08	0.34	-0.12		TENASCIN; CHAIN: A, B;	STRUCTURAL PROTEIN TENASCIN, FIBRONECTIN TYPE-III, HEPARIN, EXTRACELLULAR 2 MATRIX, ADHESION, FUSION PROTEIN, STRUCTURAL PROTEIN
1427	1ten		339	419	8e-07	-0.44	0.11		CELL ADHESION PROTEIN TENASCIN (THIRD FIBRONECTIN TYPE III REPEAT) ITEN 3	
1427	1ttn		213	298	6e-18	0.64	0.65		MUSCLE PROTEIN TITIN	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1427	1wio	A	211	364	2e-15	0.19	0.21		MODULE M5 (CONNECTIN) 1TNM3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNM 4 1TNM 58	
1427	1wit		210	298	8e-19	0.50	0.31		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
1427	1wvc	A	209	303	6e-21	0.20	0.15		TWITCHIN 18TH IGSF MODULE; CHAIN: NULL;	MUSCLE PROTEIN IMMUNOGLOBULIN SUPERFAMILY, 1 SET, MUSCLE PROTEIN
1427	1ww	X	215	302	1.4e-16	0.05	0.04		NT-3 GROWTH FACTOR RECEPTOR TRKC; CHAIN: A;	TRANSFERASE TRK RECEPTOR, RECEPTOR TYROSINE KINASE, 3D-DOMAIN SWAPPING, 2 TRANSFERASE
1427	1wtg	A	2	166	1e-16	0.19	-0.09		NERVE GROWTH FACTOR; CHAIN: V, W; TRKA RECEPTOR; CHAIN: X, Y;	NERVE GROWTH FACTOR/TRKA COMPLEX BETA-NGF; COMPLEX, TRKA RECEPTOR, NERVE GROWTH FACTOR, CYSTEINE KNOT, 2 IMMUNOGLOBULIN LIKE DOMAIN, NERVE GROWTH FACTOR/TRKA COMPLEX
1427	2bnh		1	170	4e-25	0.28	0.29		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1427									RIBONUCLEASE	ACETYLATION RNASE

836

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INHIBITOR; CHAIN: NULL;	INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1427	2bnh		2	323	1.7e-14	0.31	-0.07		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1427	2ncm		210	298	4e-16	0.29	0.70		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
1427	3ncm	A	211	298	4e-18	0.52	0.48		NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM; CHAIN: A;	CELL ADHESION PROTEIN NCAM MODULE 2; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, HOMOPHILIC 3 BINDING, CELL ADHESION PROTEIN
1437	1dan	L	897	980	1.7e-11	0.01	-0.20		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
1437	1dva	L	897	980	1.7e-11	0.06	-0.19		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX

837

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	
1437	1dx5	I	703	808	5.1e-12	0.11	-0.20		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
1437	1emn		652	729	5.1e-08	0.32	-0.19		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1437	1emn		703	767	6.8e-09	0.28	-0.19		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1437	1emn		817	884	3.4e-09	0.04	-0.19		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1437	1ezg	A	669	758	3.4e-08	0.20	-0.19		THERMAL HYSTERESIS PROTEIN ISOFORM YL-1; CHAIN: A, B;	DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILIN-1 FRAGMENT, MATRIX PROTEIN
1437	1fak	L	897	980	1.7e-11	0.16	-0.20		BLOOD COAGULATION FACTOR VIIA; CHAIN: L;	ANTIFREEZE PROTEIN INSECT
1437	1klo		509	683	1.5e-08	0.18	-0.20		GLYCOPROTEIN GLYCOPROTEIN	PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE
1437	1klo		580	735	1.7e-09	0.06	-0.20		LAMININ; CHAIN: NULL;	PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3
1437	1klo		780	921	1.7e-13	0.02	-0.20		LAMININ; CHAIN: NULL;	INHIBITOR, GLA, EGF, COMPLEX (SERINE 4
1437	1klo		824	989	3.4e-13	0.05	-0.20		LAMININ; CHAIN: NULL;	PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
1437	1klo		876	1036	3.4e-20	0.01	-0.20		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1437	1ptx	L	523	618	5.1e-09	0.11	-0.20		FACTOR IXA; CHAIN: C;	GLYCOPROTEIN GLYCOPROTEIN
1437	1qfk	L	901	980	8.5e-11	0.36	-0.17		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	
1437	1vmo	A	570	741	4.4e-31	0.49	-0.12		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I1VMO3	
1437	1vmo	A	637	800	4.4e-26	0.78	-0.15		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I1VMO3	
1437	1vmo	A	687	878	1.3e-23	0.13	-0.18		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I1VMO3	
1437	1vmo	A	721	935	8.8e-22	0.23	-0.19		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I1VMO3	
1437	1vmo	A	819	985	6.6e-26	0.34	-0.19		MEMBRANE PROTEIN VITELLINE MEMBRANE OUTER LAYER PROTEIN I1VMO3	
1437	1xka	L	777	861	1.7e-10	0.02	-0.20		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
1437	9wga	A	548	723	3.4e-12	0.27	-0.18		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	
1437	9wga	A	601	775	8.5e-15	0.02	-0.19		LECTIN (AGGLUTININ)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1437	9wga	A	647	816	8.5e-13	0.25	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1437	9wga	A	734	918	1.7e-11	-0.00	-0.20		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1437	9wga	A	765	963	1.2e-14	0.13	-0.17		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1437	9wga	A	822	990	5.1e-13	0.05	-0.18		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1437	9wga	A	848	1016	5.1e-15	0.22	-0.17		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1442	1a4y	A	56	242	1.5e-15	0.11	0.28		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1442	1a9n	A	59	145	4.4e-09	-0.04	0.47		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B';	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA,



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1442	1a9n	A	66	187	1.5e-22	0.46	1.00		CHAIN: B, D; U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	SNRNP, RIBONUCLEOPROTEIN COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1442	1a9n	C	59	158	6.6e-10	0.18	0.63		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1442	1a9n	C	66	187	6.6e-22	0.17	0.98		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1442	1a9n	C	89	200	1.8e-16	0.43	0.16		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1442	1d0b	A	36	163	5.1e-21	0.67	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1442	1d0b	A	39	186	8.8e-15	0.23	-0.07		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1442	1dce	A	38	140	3.4e-10	0.73	0.89		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1442	1dce	A	57	163	6.8e-12	0.12	0.93		RAB GERANYLGERANYLTRANSFERASE ALPHA	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	E, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1442	1ds9	A	47	162	1.7e-11	-0.02	0.84		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1442	1fgv	A	35	162	6.8e-08	0.42	0.17		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1442	1fs2	A	35	162	6.8e-08	0.23	0.22		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1442	1yr8	A	59	186	8.8e-14	-0.06	0.63		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAI; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAI, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMISPHERAL TWINNING, 3 MEROPHEDRAL TWINNING, MEROPHEDRY
1442	2bnh		56	266	1.3e-16	0.06	-0.05		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1442	2bnh		59	192	6.6e-20	-0.05	0.03		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1443	1aab		2	81	1e-27			73.20	HIGH MOBILITY GROUP PROTEIN; 1AAB 5 CHAIN: NULL; 1AAB 6	DNA-BINDING HMG A DNA-BINDING HMG-BOX DOMAIN A OF RAT HMG1; 1AAB 8 HMG-BOX 1AAB 20
1443	1aab		3	75	1e-27	0.34	1.00		HIGH MOBILITY GROUP PROTEIN; 1AAB 5 CHAIN: NULL; 1AAB 6	DNA-BINDING HMG A DNA-BINDING HMG-BOX DOMAIN A OF RAT HMG1; 1AAB 8 HMG-BOX 1AAB 20
1443	1aab		83	123	1.5e-07	-0.51	0.17		HIGH MOBILITY GROUP PROTEIN; 1AAB 5 CHAIN: NULL; 1AAB 6	DNA-BINDING HMG A DNA-BINDING HMG-BOX DOMAIN A OF RAT HMG1; 1AAB 8 HMG-BOX 1AAB 20
1443	1cct	A	5	75	6.8e-27	0.11	1.00		HIGH MOBILITY GROUP 1 PROTEIN; CHAIN: A; DNA (5'-D(*CP*CP*(IDO) CHAIN: B; DNA (5'-CHAIN: C;	GENE REGULATION/DNA HMG-1, AMPHOTERIN, HEPARIN-BINDING PROTEIN P30; HIGH-MOBILITY GROUP DOMAIN, BENT DNA, PROTEIN-DRUG-DNA 2 COMPLEX, GENE REGULATION/DNA
1443	1cct	A	5	75	6.8e-27			65.26	HIGH MOBILITY GROUP 1 PROTEIN; CHAIN: A; DNA (5'-D(*CP*CP*(IDO) CHAIN: B; DNA (5'-CHAIN: C;	GENE REGULATION/DNA HMG-1, AMPHOTERIN, HEPARIN-BINDING PROTEIN P30; HIGH-MOBILITY GROUP DOMAIN, BENT DNA, PROTEIN-DRUG-DNA 2 COMPLEX, GENE REGULATION/DNA
1443	1hne		91	159	3.4e-10	-0.49	0.11		DNA-BINDING HIGH MOBILITY GROUP PROTEIN FRAGMENT-B	

Table 5

SEQ ID NO:	PDB ID	CHAI NID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(HMG1) (DNA-BINDING 1HME 3 HMG-BOX DOMAIN B OF RAT HMG1) (NMR, 1 STRUCTURE) 1HME 4	
1443	1hsm		91	163	5.1e-11	-0.42	0.42		DNA-BINDING HIGH MOBILITY GROUP PROTEIN 1 (HMG1) BOX 2, COMPLEXED WITH 1HSM 3	
									MERCAPTOETHANOL (NMR, MINIMIZED AVERAGE STRUCTURE) 1HSM 4	
1443	1qrv	A	88	123	1.1e-08	-0.86	0.22		DNA (5'-D(*GP*CP*GP*AP*TP*AP*TP*CP*GP*C)-3'); CHAIN: C, D; HIGH MOBILITY GROUP PROTEIN D; CHAIN: A, B;	GENE REGULATION/DNA HMG-D; PROTEIN-DNA COMPLEX, HMG DOMAIN, NON-SEQUENCE SPECIFIC 2 CHROMOSOMAL PROTEIN HMG-D
1477	1mey	C	232	316	1.4e-35	-0.63	0.06		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1477	1ubd	C	214	316	9.8e-29	-0.76	0.16		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1482	1ctq	A	447	482	1.2e-06	-0.54	0.10		TRANSFORMING PROTEIN P21/H-RAS-1;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A;	CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
1482	1dpt	A	400	482	6e-07	0.05	0.10		RHOA; CHAIN: A;	GENE REGULATION/SIGNALING PROTEIN RHOA-GDP COMPLEX
1482	1huq	A	443	482	3.6e-07	-0.88	0.41		RAB5C; CHAIN: A;	PROTEIN TRANSPORT G-PROTEIN, GTP HYDROLYSIS, ENDOCYTOSIS, RAB PROTEIN, 2 MEMBRANE TRAFFICKING
1484	1buo	A	71	190	8.4e-16	-0.01	0.21		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
1487	1mm1		395	507	0.0018	-0.46	0.09		MM1V REVERSE TRANSCRIPTASE; 1MML 4 CHAIN: NULL; 1MML 5	REVERSE TRANSCRIPTASE
1489	1fgx	A	104	403	0	0.67	1.00		BETA 1,4 GALACTOSYLTRANSFERASE; CHAIN: A, B;	TRANSFERASE BETA4GALT1; NUCLEOTIDE BINDING PROTEIN, ALPHA BETA ALPHA FOLD
1489	1fgx	A	94	404	0			343.91	BETA 1,4 GALACTOSYLTRANSFERASE; CHAIN: A, B;	TRANSFERASE BETA4GALT1; NUCLEOTIDE BINDING PROTEIN, ALPHA BETA ALPHA FOLD
1491	1c1g	A	22	301	4.2e-10			61.55	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
1491	1dgs	A	23	145	7e-31	-0.19	0.90		INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1; CHAIN: A;	SIGNALING PROTEIN GUANINE NUCLEOTIDE-BINDING PROTEIN 1; GBP, GTP HYDROLYSIS, GDP,

846

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1491	1f5n	A	23	145	7e-31	-0.39	0.92		INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1; CHAIN: A;	GMP, INTERFERON INDUCED, DYNAMIN 2 RELATED, LARGE GTPASE FAMILY, SIGNALING PROTEIN
1491	1i1i	P	25	273	7.2e-09	-0.34	0.01		NEUROLYSIN; CHAIN: P;	HYDROLASE NEUROPEPTIDASE, ZINC METALLOPEPTIDASE, ENDOPEPTIDASE
1493	1g0s	A	297	396	3.6e-07	-0.10	0.00		HYPOTHETICAL 23.7 KDA PROTEIN IN ICC-TOLC CHAIN: A, B;	HYDROLASE ADP-RIBOSE PYROPHOSPHATASE; NUDIX FOLD
1497	1b0w	A	1	95	7e-57			113.54	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C;	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
1497	1b0w	A	23	125	2.4e-63			123.78	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C;	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
1497	1b6d	A	23	124	5.6e-65	0.86	1.00		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN
1497	1b1l	J	23	124	2.8e-66	0.85	1.00		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
1497	1bvk	A	1	95	2.8e-56			113.70	HULYS11; CHAIN: A, B, D, E, LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED

847

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1497	1bvk	A	23	119	1.2e-63	0.95	1.00		HULYS11; CHAIN: A, B, D, E, LYSOZYME; CHAIN: C, F;	ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
1497	1bvk	A	23	125	1.2e-63			123.99	HULYS11; CHAIN: A, B, D, E, LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
1497	1bw	A	1	95	8.4e-59			116.30	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENGE-JONES 2 PROTEIN, IMMUNE SYSTEM
1497	1bw	A	21	125	2.8e-65			127.42	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENGE-JONES 2 PROTEIN, IMMUNE SYSTEM
1497	1bw	A	23	124	2.8e-65	0.85	1.00		IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENGE-JONES 2 PROTEIN, IMMUNE SYSTEM
1497	1dee	A	23	124	7e-68	1.07	1.00		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1497	1dql	L	1	95	5.6e-57			111.65	CHAIN: G, H; IGM MEZ IMMUNOGLOBULIN; CHAIN: L; IGM MEZ IMMUNOGLOBULIN; CHAIN: H;	FAB VH3 3 SPECIFICITY IMMUNE SYSTEM IMMUNOGLOBULIN FOLD, ANTIBODY, IGM, FV
1497	1dql	L	23	125	7e-63			123.32	IGM MEZ IMMUNOGLOBULIN; CHAIN: L; IGM MEZ IMMUNOGLOBULIN; CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD, ANTIBODY, IGM, FV
1497	1fgv	L	1	93	1.4e-60			121.32	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	
1497	1fgv	L	23	124	4.2e-66	0.97	1.00		IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	
1497	1fgv	L	23	125	4.2e-66			131.86	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	
1497	1fvc	A	1	95	1.1e-57			115.15	IMMUNOGLOBULIN FV FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 8 1FVC3	
1497	1fvc	A	23	124	5.6e-64	0.82	1.00		IMMUNOGLOBULIN FV	



Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 8 IFVC 3	
1497	1fvc	A	23	125	5.6e-64			125.91	IMMUNOGLOBULIN FV FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 8 IFVC 3	
									IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 IFVD 3	
1497	1fvd	A	23	124	2.8e-64	1.00	1.00		IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
								113.29	IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
1497	1igm	L	1	93	8.4e-59				IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
									IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
1497	1igm	L	23	124	4.2e-64	0.86	1.00		IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
									IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
1497	1igm	L	23	125	4.2e-64			123.98	IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
									IMMUNOGLOBULIN M (IG-M) FV FRAGMENT IIGM 3	
1497	1wtl	A	1	93	4.2e-57			116.77	IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN IWTL 3 (BENCE-JONES PROTEIN) IWTL 4	
									IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN IWTL 3 (BENCE-JONES PROTEIN) IWTL 4	
1497	1wtl	A	23	125	1.2e-63			127.46	IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									LIGHT-CHAIN 1 WTL 3 (BENCE-JONES PROTEIN) 1 WTL 4	
1497	2fgw	L	23	124	7e-67	0.94	1.00		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
1498	1a80		104	367	5.6e-69	0.12	0.99		2,5-DIKETO-D-GLUCONIC ACID REDUCTASE A; CHAIN: NULL;	OXIDOREDUCTASE 2,5-DKG REDUCTASE A; OXIDOREDUCTASE, ALPHA8/BETA8 BARREL, 2,5-DIKETO-D-GLUCONIC 2 ACID, COMMERCIAL VITAMIN C SYNTHESIS
1498	1a80		122	372	5.6e-69			62.58	2,5-DIKETO-D-GLUCONIC ACID REDUCTASE A; CHAIN: NULL;	OXIDOREDUCTASE 2,5-DKG REDUCTASE A; OXIDOREDUCTASE, ALPHA8/BETA8 BARREL, 2,5-DIKETO-D-GLUCONIC 2 ACID, COMMERCIAL VITAMIN C SYNTHESIS
1498	1ads		105	365	2.8e-84	0.05	0.96		OXIDOREDUCTASE ALDOSE REDUCTASE (E.C.1.1.1.21) COMPLEX WITH NADPH 1ADS 3	
1498	1ads		82	372	2.8e-84			105.14	OXIDOREDUCTASE ALDOSE REDUCTASE (E.C.1.1.1.21) COMPLEX WITH NADPH 1ADS 3	
1498	1afs	A	68	372	1.1e-72			101.13	3-ALPHA-HYDROXYSTEROID DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE 3-ALPHA-HSD; OXIDOREDUCTASE, NAD

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1498	1af5	A	99	368	1.1e-72	0.16	0.83		3-ALPHA-HYDROXYSTEROID DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE 3-ALPHA-HSD; OXIDOREDUCTASE, NAD
1498	1ah4		105	365	1.1e-83	-0.09	0.94		ALDOSE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, ALDOSE REDUCTASE, INHIBITION, DIABETES
1498	1ah4		81	372	1.1e-83			103.47	ALDOSE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, ALDOSE REDUCTASE, INHIBITION, DIABETES
1498	1c9w	A	105	365	2.8e-81	0.12	0.84		CHO REDUCTASE; CHAIN: A;	OXIDOREDUCTASE ALPHA/BETA TIM BARREL, PROTEIN-NADP+ COMPLEX
1498	1c9w	A	80	372	2.8e-81			103.01	CHO REDUCTASE; CHAIN: A;	OXIDOREDUCTASE ALPHA/BETA TIM BARREL, PROTEIN-NADP+ COMPLEX
1498	1cwn		106	368	9.8e-78	0.09	0.70		ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; TIM-BARREL, OXIDOREDUCTASE, NADP
1498	1cwn		79	372	9.8e-78			81.73	ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; TIM-BARREL, OXIDOREDUCTASE, NADP
1498	1ei3	A	102	365	2.8e-85	0.21	0.99		ALDOSE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE ALDOSE REDUCTASE, INHIBITION, DIABETES
1498	1ei3	A	82	372	2.8e-85			106.54	ALDOSE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE ALDOSE REDUCTASE, INHIBITION, DIABETES
1498	1fb		105	365	1.4e-80	0.08	0.90		FR-1 PROTEIN; CHAIN: NULL;	OXIDOREDUCTASE (NADP) ALDO-KETO OXIDOREDUCTASE (NADP), TIM BARREL
1498	1fb		80	372	1.4e-80			98.60	FR-1 PROTEIN; CHAIN: NULL;	OXIDOREDUCTASE (NADP)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									NULL;	ALDO-KETO OXIDOREDUCTASE (NADP), TIM BARREL
1498	1hw6	A	104	367	1.4e-66	0.15	0.90		2,5-DIKETO-D-GLUCONIC ACID REDUCTASE; CHAIN: A;	OXIDOREDUCTASE APO-2,5-DIKETO-D-GLUCONATE REDUCTASE; ALDO-KETO REDUCTASE, TIM BARREL
1498	1hw6	A	113	365	1.4e-66			67.97	2,5-DIKETO-D-GLUCONIC ACID REDUCTASE; CHAIN: A;	OXIDOREDUCTASE APO-2,5-DIKETO-D-GLUCONATE REDUCTASE; ALDO-KETO REDUCTASE, TIM BARREL
1498	2air		106	368	1.1e-76	0.22	0.70		ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; OXIDOREDUCTASE, TIM-BARREL
1498	2air		79	372	1.1e-76			61.33	ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; OXIDOREDUCTASE, TIM-BARREL
1499	1flm	A	543	626	1.2e-18	0.33	-0.18		OUTER SURFACE PROTEIN C; CHAIN: A, B, C, D;	IMMUNE SYSTEM OSPC; LYME DISEASE, OSPC, HB19, HELICAL BUNDLE, DIMER
1499	1flm	A	550	626	3.6e-17	0.14	-0.15		OUTER SURFACE PROTEIN C; CHAIN: A, B, C, D;	IMMUNE SYSTEM OSPC; LYME DISEASE, OSPC, HB19, HELICAL BUNDLE, DIMER
1499	1flm	A	551	626	9.6e-17	0.32	-0.20		OUTER SURFACE PROTEIN C; CHAIN: A, B, C, D;	IMMUNE SYSTEM OSPC; LYME DISEASE, OSPC, HB19, HELICAL BUNDLE, DIMER
1499	1g5z	A	550	626	1.2e-17	0.36	-0.19		OUTER SURFACE PROTEIN C; CHAIN: A;	IMMUNE SYSTEM SURFACE PROTEIN, ALPHA HELIX PROTEIN
1499	1g5z	A	550	626	3.6e-17	0.04	-0.20		OUTER SURFACE PROTEIN C; CHAIN: A;	IMMUNE SYSTEM SURFACE PROTEIN, ALPHA HELIX PROTEIN
1499	1g5z	A	551	626	1.2e-17	0.34	-0.19		OUTER SURFACE PROTEIN C; CHAIN: A;	IMMUNE SYSTEM SURFACE PROTEIN, ALPHA HELIX PROTEIN
1499	1g5z	A	551	626	1.2e-17	0.52	-0.20		OUTER SURFACE PROTEIN C; CHAIN: A;	IMMUNE SYSTEM SURFACE PROTEIN, ALPHA HELIX PROTEIN
1499	1g5z	A	551	626	2.4e-17	0.18	-0.19		OUTER SURFACE	IMMUNE SYSTEM SURFACE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1499	1g5z	A	551	626	3.6e-17	0.67	-0.19		PROTEIN C; CHAIN: A; OUTER SURFACE PROTEIN C; CHAIN: A;	PROTEIN, ALPHA HELIX PROTEIN IMMUNE SYSTEM SURFACE PROTEIN, ALPHA HELIX PROTEIN
1504	1ckl	A	10	79	1.4e-30			90.17	CD46; CHAIN: A, B, C, D, E, F;	GLYCOPROTEIN MEMBRANE COFACTOR PROTEIN (MCP); VIRUS RECEPTOR, COMPLEMENT COFACTOR, SHORT CONSENSUS REPEAT, 2 SCR, MEASLES VIRUS, GLYCOPROTEIN
1504	1ckl	A	1	72	2.8e-24			89.35	CD46; CHAIN: A, B, C, D, E, F;	GLYCOPROTEIN MEMBRANE COFACTOR PROTEIN (MCP); VIRUS RECEPTOR, COMPLEMENT COFACTOR, SHORT CONSENSUS REPEAT, 2 SCR, MEASLES VIRUS, GLYCOPROTEIN
1504	1ckl	A	222	347	1.4e-27			222.13	CD46; CHAIN: A, B, C, D, E, F;	GLYCOPROTEIN MEMBRANE COFACTOR PROTEIN (MCP); VIRUS RECEPTOR, COMPLEMENT COFACTOR, SHORT CONSENSUS REPEAT, 2 SCR, MEASLES VIRUS, GLYCOPROTEIN
1504	1g40	A	77	344	7e-35			81.51	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B;	IMMUNE SYSTEM BETA, MODULE
1504	1g44	A	166	406	5.6e-31			69.90	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1504	1g44	B	220	406	1.4e-25			74.78	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1504	1g44	C	153	406	5.6e-24			69.94	COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1504	1qub	A	74	406	4.2e-28			74.76	HUMAN BETA2-	MEMBRANE ADHESION SHORT

854

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GLYCOPROTEIN I; CHAIN: A;	CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
1511	1alh	A	281	363	4.2e-29	-0.50	0.03		QCSR ZINC FINGER PEPTIDE; CHAIN: A; DUPELEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1511	1alh	A	367	436	2.8e-26	-0.33	0.28		QCSR ZINC FINGER PEPTIDE; CHAIN: A; DUPELEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1511	1bbo		369	424	4.2e-14	-0.42	0.42		DNA-BINDING PROTEIN HUMAN ENHANCER- BINDING PROTEIN MBP- 1 MUTANT WITH CYS 11 1BBO 3 REPLACED BY ABU (C11ABU) (NMR, 60 STRUCTURES) 1BBO 4	
1511	1mey	C	10	73	1.4e-40			64.91	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1511	1mey	C	280	363	2.8e-47	-0.07	0.39		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1511	1mey	C	309	391	2.8e-49	-0.12	1.00		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA)

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1511	1mey	C	338	421	2.8e-49	0.17	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1511	1mey	C	338	424	2.8e-49			83.28	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1511	1mey	C	366	437	2.8e-40	-0.25	0.15		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1511	1mey	G	336	363	5.6e-13	0.73	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1511	1t66	A	235	400	1.4e-35	-0.57	0.33		TFIIIA; CHAIN: A, D: 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1511	1t66	A	279	449	1.4e-35			65.99	TFIIIA; CHAIN: A, D: 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1511	1tf6	A	339	447	4.2e-24	-0.51	0.06		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1511	1ubd	C	309	420	6e-35			68.77	YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1511	1ubd	C	314	419	1.4e-34	-0.02	0.87		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1511	1ubd	C	324	403	6e-35	-0.34	0.37		YY1; CHAIN: C; ADENO- ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1511	1ubd	C	346	436	4.2e-28	-0.08	0.15		YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION REGULATION/DNA)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B,	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1511	2gi	A	288	421	2.8e-33	-0.30	0.22		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1511	2gi	A	309	449	2.8e-33			63.97	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1511	2gi	A	317	435	7e-31	0.05	0.77		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1511	2gi	A	346	446	1.1e-22	0.09	-0.15		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1513	1hi7	A	139	173	0.0015	-0.26	0.06		PS2 PROTEIN; CHAIN: A, B;	GROWTH FACTOR PNR-2, PS2, TFF1, BREAST CANCER ESTROGEN INDUCIBLE GROWTH FACTOR, CELL MOTILITY, TUMOR SUPPRESSOR, TREFOIL 2 DOMAIN, SIGNAL
1518	1bmng		21	116	1.4e-34			131.75	BETA=2=, MICROGLOBULIN; 1BMG 5 CHAIN: NULL 1BMG 6	HISTOCOMPATIBILITY ANTIGEN LACTOLIN, MHC-1 HISTOCOMPATIBILITY ANTIGEN, LIGHT CHAIN 1BMG 11
1518	1i4f	B	1	85	7e-31			134.89	HLA CLASS I	IMMUNE SYSTEM MAGE-4

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1518	1i4f	B	20	116	2.8e-34			157.82	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM IMAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1521	1bmp		12	115	4.2e-49			156.36	BONE MORPHOGENETIC PROTEIN-7; CHAIN: NULL;	TRANSFORMING GROWTH FACTOR OSTEOGENIC PROTEIN-1, HOP-1, BMP-7; MORPHOGEN, TRANSFORMING GROWTH FACTOR, CYTOKINE, BONE, 2 CARTILAGE, GLYCOPROTEIN
1521	1bmp		638	741	4.2e-50	0.36	1.00		BONE MORPHOGENETIC PROTEIN-7; CHAIN: NULL;	TRANSFORMING GROWTH FACTOR OSTEOGENIC PROTEIN-1, HOP-1, BMP-7; MORPHOGEN, TRANSFORMING GROWTH FACTOR, CYTOKINE, BONE, 2 CARTILAGE, GLYCOPROTEIN
1521	1bmp		638	741	4.2e-50			156.52	BONE MORPHOGENETIC PROTEIN-7; CHAIN: NULL;	TRANSFORMING GROWTH FACTOR OSTEOGENIC PROTEIN-1, HOP-1, BMP-7; MORPHOGEN, TRANSFORMING GROWTH FACTOR, CYTOKINE, BONE, 2 CARTILAGE, GLYCOPROTEIN
1521	1poi	A	40	257	7e-39	0.29	0.62		GLUTACONATE COENZYM A-	TRANSFERASE TRANSFERASE, COA, GLUTAMATE, PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1521	1poi	A	53	244	1.1e-51	0.22	0.21		TRANSFERASE; CHAIN: A, B, C, D;	FERMENTATION
									GLUTACONATE COENZYME A-TRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE TRANSFERASE, COA, GLUTAMATE, PROTEIN FERMENTATION
1521	1poi	B	270	490	3.6e-60	0.13	0.93		GLUTACONATE COENZYME A-TRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE TRANSFERASE, COA, GLUTAMATE, PROTEIN FERMENTATION
1521	1poi	B	273	501	2.8e-34	0.01	0.83		GLUTACONATE COENZYME A-TRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE TRANSFERASE, COA, GLUTAMATE, PROTEIN FERMENTATION
1523	1atl	A	221	422	4.2e-39	0.70	1.00		ATROLYSIN C; 1ATL 4 CHAIN: A, B, C, D; 1ATL 5	METALLOENDOPEPTIDASE HEMORRHAGIC TOXIN C, FORM D; 1ATL 6
1523	1atl	A	223	422	1.2e-42	0.73	1.00		ATROLYSIN C; 1ATL 4 CHAIN: A, B, C, D; 1ATL 5	METALLOENDOPEPTIDASE HEMORRHAGIC TOXIN C, FORM D; 1ATL 6
1523	1bkc	A	222	421	1.2e-43	-0.25	0.17		TUMOR NECROSIS FACTOR-ALPHA-CONVERTING ENZYME; CHAIN: A, C, E, I;	ZN-ENDOPEPTIDASE TACE; ZN-ENDOPEPTIDASE, HYDROLASE, TNF-ALPHA
1523	1bkc	A	225	417	9.8e-07	0.03	0.01		TUMOR NECROSIS FACTOR-ALPHA-CONVERTING ENZYME; CHAIN: A, C, E, I;	ZN-ENDOPEPTIDASE TACE; ZN-ENDOPEPTIDASE, HYDROLASE, TNF-ALPHA
1523	1bud	A	222	420	2.8e-36	0.65	0.87		ACUTOLYSIN A; CHAIN: A;	TOXIN HEMORRHAGIN I, 1AAH-I; METALLOPROTEINASE, SNAKE VENOM, MMP, TOXIN
1523	1dan	L	662	718	1.4e-07	0.34	0.23		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFRCMK) WITH CHAIN: C;	INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
1523	1dva	L	662	718	1.4e-07	0.53	0.18		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
1523	1emm		589	656	2.8e-11	0.15	-0.17		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1523	1emm		629	690	1.4e-09	0.19	0.06		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1523	1emm		663	718	4.2e-07	0.40	-0.01		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1523	1fak	L	662	718	1.4e-07	0.64	0.25		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	DOMAIN, HUMAN FIBRILIN-1 FRAGMENT, MATRIX PROTEIN
1523	1fvl		439	508	4.2e-15	-0.04	0.84		FLAVORIDIN; 1FVL 4 CHAIN: NULL 1FVL 5	BLOOD COAGULATION INHIBITOR GP IIB/IIIA ANTAGONIST 1FVL 9
1523	1fvl		439	512	1.2e-26	0.36	0.94		FLAVORIDIN; 1FVL 4 CHAIN: NULL 1FVL 5	BLOOD COAGULATION INHIBITOR GP IIB/IIIA ANTAGONIST 1FVL 9
1523	1iag		221	422	2.8e-38	0.77	1.00		METALLOPROTEINASE ADAMALYSIN II (PROTEINASE II) (E.C.3.4.24.46) IIAG 3	
1523	1iag		223	422	4.8e-41	0.65	1.00		METALLOPROTEINASE ADAMALYSIN II (PROTEINASE II) (E.C.3.4.24.46) IIAG 3	
1523	1igr	A	410	706	1.2e-11	0.15	-0.19		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
1523	1kdo		461	630	7e-20	0.09	-0.13		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1523	1kdo		519	664	2.8e-11	0.22	0.17		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1523	1kdo		638	774	1.4e-10	0.07	-0.20		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1523	1kst		439	508	1.3e-15	0.18	0.35		AGGREGATION INHIBITOR, GP	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ANTAGONIST KISTRIN (NMR, 8 STRUCTURES) IKST 3	
1523	Ikst		439	510	3.6e-25	0.31	0.60		AGGREGATION INHIBITOR, GP ANTAGONIST KISTRIN (NMR, 8 STRUCTURES) IKST 3	
1523	Ipfx	L	662	718	9.8e-08	0.53	0.72		FACTOR IXA; CHAIN: C <sub>1</sub> , D-PHE-PRO-ARG; CHAIN: I <sub>1</sub>	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR: COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
1523	Iqua	A	221	420	3.6e-39	0.63	1.00		ACUTOLYSIN-C; CHAIN: A <sub>1</sub>	TOXIN HEMORRHAGIN III METALLOPROTEASE, HEMORRHAGIC TOXIN, SNAKE VENOM/ PROTEINASE, 2 CRYSTAL STRUCTURE, AGKISTRODON ACUTUS
1523	Iqua	A	221	420	9.8e-35	0.82	1.00		ACUTOLYSIN-C; CHAIN: A <sub>1</sub>	TOXIN HEMORRHAGIN III METALLOPROTEASE, HEMORRHAGIC TOXIN, SNAKE VENOM/ PROTEINASE, 2 CRYSTAL STRUCTURE, AGKISTRODON ACUTUS
1523	Iqub	A	414	693	1.2e-11	0.01	-0.17		HUMAN BETA2-GLYCOPROTEIN I; CHAIN: A <sub>1</sub>	MEMBRANE ADHESION SHORT CONSENSUS REPEAT, SUSHI, COMPLEMENT CONTROL PROTEIN, 2 N-GLYCOSYLATION, MULTI-DOMAIN, MEMBRANE ADHESION
1523	Zech		469	516	2.4e-18	-0.19	0.30		BLOOD COAGULATION	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INHIBITOR ECHISTATIN (NMR, 8 STRUCTURES) 2ECH 3	
1523	9wga	A	502	666	2.8e-16	0.27	0.65		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1523	9wga	A	526	701	1.4e-12	0.09	0.23		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1523	9wga	A	661	859	7e-15	0.08	-0.14		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1524	1alh	A	21	103	5.6e-31			75.55	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1524	1alh	A	2	84	4.2e-31			77.96	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1524	1alh	A	58	140	4.2e-31			77.62	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1524	1atl	A	8	210	1.4e-74			104.44	ATROLYSIN C; IATL 4	METALLOENDOPEPTIDASE

864

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL ID score	Compound	PDB annotation
									CHAIN: A, B, C, D; 1ATL 5	HEMORRHAGIC TOXIN C, FORM D; 1ATL 6
1524	1bud	A	10	208	5.6e-71			97.67	ACUTOLYSIN A; CHAIN: A;	TOXIN HEMORRHAGIN 1, 1AAH-I; METALLOPROTEINASE, SNAKE VENOM, MMP, TOXIN
1524	1efe		30	184	4.2e-42			82.74	PATHOGENESIS-RELATED PROTEIN P14A; CHAIN: NULL;	PATHOGENESIS-RELATED PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
1524	1fvl		13	88	1.4e-25			67.46	FLAVORIDIN; 1FVL 4 CHAIN: NULL, 1FVL 5	BLOOD COAGULATION INHIBITOR GP IIB/IIIA ANTAGONIST 1FVL 9
1524	1iag		7	210	4.2e-74			105.32	METALLOPROTEASE ADAMALYSIN II (PROTEINASE II) (E.C.3.4.24.46) 1IAG 3	
1524	1jln	A	21	247	2.8e-45	-0.00	-0.09		PROTEIN TYROSINE PHOSPHATASE, RECEPTOR TYPE, R; CHAIN: A;	HYDROLASE STEP-LIKE PTPASE, PROTEIN-TYROSINE-PHOSPHATASE PROTEIN TYROSINE PHOSPHATASE, PTP-SL, PTPBR7, ERK2-MAP 2 KINASE REGULATION
1524	1kst		13	85	4.2e-25			66.02	AGGREGATION INHIBITOR, GP ANTAGONIST KISTRIN (NMR, 8 STRUCTURES) 1KST 3	
1524	1mcy	C	1	64	2.8e-38			60.19	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1524	1mey	C	1	66	1.1e-39			68.19	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1524	1mey	C	29	111	4.2e-50			93.01	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1524	1mey	C	48	130	2.8e-50			95.82	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1524	1mey	C	57	139	1.4e-50			95.58	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1524	1mey	C	57	139	2.8e-50			95.19	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1524	1mkp		100	242	1.2e-31			132.82	PYSTI; CHAIN: NULL;	HYDROLASE DUAL SPECIFICITY PHOSPHATASE, MAP KINASE HYDROLASE
1524	1mkp		9	151	2.8e-34			133.40	PYSTI; CHAIN: NULL;	HYDROLASE DUAL SPECIFICITY PHOSPHATASE, MAP KINASE HYDROLASE
1524	1qnx	A	1	187	4.2e-44			81.43	VES V 5; CHAIN: A;	ALLERGEN ANTIGEN 5; ANTIGEN 5, ALLERGEN, VESPID VENOM
1524	1qua	A	7	208	7e-69			109.26	ACUTOLYSIN-C; CHAIN:	TOXIN HEMORRHAGIN III

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A;	METALLOPROTEASE, HEMORRHAGIC TOXIN, SNAKE VENOM PROTEINASE, 2 CRYSTAL STRUCTURE, AGKISTRODON ACUTUS
1524	Irpm	A	18	241	2.8e-59	-0.04	0.35		RECEPTOR PROTEIN TYROSINE PHOSPHATASE MU; CHAIN: A, B;	RECEPTOR D1; RECEPTOR, PHOSPHATASE, SIGNAL TRANSDUCTION, ADHESION, 2 HYDROLASE
1524	1tt6	A	1	153	4.2e-37			71.61	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1524	1tt6	A	1	153	5.6e-37			107.29	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1524	1tt6	A	1	153	7e-38			101.40	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1524	1tt6	A	3	144	8.4e-34			66.54	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1524	1ubd	C	1	111	4.2e-36			84.52	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
1524	1ubd	C	20	130	5.6e-36			88.82	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1524	1ubd	C	31	139	2.8e-36			88.27	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1524	1ubd	C	3	111	1.4e-35			83.28	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1524	1vhr	A	81	250	2.4e-34			87.98	HUMAN VHI-RELATED	HYDROLASE VHR; HYDROLASE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DUAL-SPECIFICITY PHOSPHATASE CHAIN: A, B;	PROTEIN DUAL-SPECIFICITY PHOSPHATASE
1524	1vhr	A	88	249	2.4e-34	0.36	1.00		HUMAN VHL-RELATED DUAL-SPECIFICITY PHOSPHATASE CHAIN: A, B;	HYDROLASE VHR; HYDROLASE, PROTEIN DUAL-SPECIFICITY PHOSPHATASE
1524	2gli	A	1	131	7e-33			76.01	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1524	2gli	A	1	140	2.8e-33			83.07	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1524	2gli	A	1	140	2.8e-33			88.13	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1524	2gli	A	1	140	7e-34			91.46	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1527	1dus	A	321	387	0.0036	0.18	0.34		MJ0882; CHAIN: A;	STRUCTURAL GENOMICS HYPOTHETICAL PROTEIN, METHANOCOCCUS JANNASCHII
1527	1g6q	1	293	382	0.00048	0.14	0.18		HNKRP ARGININE N-METHYLTRANSFERASE; CHAIN: 1, 2, 3, 4, 5, 6;	TRANSFERASE SAM-BINDING DOMAIN, BETA-BARREL, MIXED ALPHA-BETA, HEXAMER, 2 DIMER
1528	1aj4		52	207	5.6e-43	-0.34	0.15		TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC, CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1528	1au1	B	245	399	1.3e-34	-0.11	0.00		SERINE/THREONINE PHOSPHATASE 2B; CHAIN: A, B;	HYDROLASE CALCINEURIN; HYDROLASE, PHOSPHATASE, IMMUNOSUPPRESSION
1528	1cdm	A	251	391	4.2e-53	-0.23	0.29		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
1528	1cdm	A	324	415	4.2e-28	-0.18	0.11		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
1528	1cll		251	391	5.6e-58	-0.39	0.40		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CELL 3	
1528	1cll		324	415	7e-29	0.03	0.11		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CELL 3	
1528	1cmf		322	391	1.4e-30	-0.09	0.59		CALMODULIN (VERTEBRATE); 1CMF 6 CHAIN: NULL; 1CMF 7	CALCIUM-BINDING PROTEIN CALMODULIN APO TR2C-DOMAIN; 1CMF 9
1528	1cll	A	52	207	1.1e-39	-0.09	0.23		CARDIAC TROPONIN C; CHAIN: A;	STRUCTURAL PROTEIN HELIX-TURN-HELIX
1528	1exr	A	249	390	5.6e-56	0.10	0.05		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
1528	1exr	A	320	415	9.8e-28	-0.09	0.16		CALMODULIN; CHAIN:	METAL TRANSPORT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A <sub>1</sub>	CALMODULIN, HIGH RESOLUTION, DISORDER
1528	1fw4	A	255	310	2.8e-15	0.09	-0.12		CALMODULIN; CHAIN: A <sub>1</sub>	METAL BINDING PROTEIN EF-HAND, HELIX-LOOP-HELIX, FRAGMENT, CALCIUM, TR2C, C-2 TERMINAL DOMAIN, CALMODULIN
1528	1fw4	A	327	390	9.8e-30	0.25	0.65		CALMODULIN; CHAIN: A <sub>1</sub>	METAL BINDING PROTEIN EF-HAND, HELIX-LOOP-HELIX, FRAGMENT, CALCIUM, TR2C, C-2 TERMINAL DOMAIN, CALMODULIN
1528	1g8i	A	234	386	2.8e-26	-0.23	0.07		NEURONAL CALCIUM SENSOR 1; CHAIN: A, B, C	METAL BINDING PROTEIN FREQUENIN; CALCIUM BINDING-PROTEIN, EF-HAND, CALCIUM ION
1528	1jfo	A	63	203	7e-17	-0.56	0.05		OBELIN; CHAIN: A <sub>1</sub>	LUMINESCENT PROTEIN BIOLUMINESCENCE, CALCIUM-REGULATED PHOTOPROTEIN, OBELIN, 2 OBELIA, HYDROID
1528	1jf2	A	63	203	1.1e-16	-0.09	0.12		OBELIN; CHAIN: A <sub>1</sub>	LUMINESCENT PROTEIN BIOLUMINESCENCE, CALCIUM-REGULATED PHOTOPROTEIN, OBELIN, 2 OBELIA, HYDROID
1528	1tcf		245	391	4.2e-46	-0.14	0.51		TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
1528	1tnx		251	390	4.2e-43	0.04	0.10		TROPONIN C; ITNX 4 CHAIN: NULL; ITNX 5	CALCIUM-BINDING PROTEIN EF-HAND ITNX 14

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1528	ltmx		324	415	9.8e-26	-0.22	0.07		TROPONIN C; ITNX 4 CHAIN: NULL; ITNX 5	CALCIUM-BINDING PROTEIN EF-HAND ITNX 14
1528	ltop		251	393	9.8e-47	-0.07	0.11		CONTRACTILE SYSTEM PROTEIN TROPONIN C ITOP 3	
1528	ltop		324	415	9.8e-26	-0.41	0.15		CONTRACTILE SYSTEM PROTEIN TROPONIN C ITOP 3	
1528	ltvk	A	249	390	2.8e-56	-0.01	0.22		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
1529	ldb3	A	153	449	4.2e-61	-0.31	0.05		GDP-MANNOSE 4,6-DEHYDRATASE; CHAIN: A;	LYASE DEHYDRATASE, NADP, GDP-MANNOSE, GDP-FUCOSE
1533	la17		39	185	2.8e-13	0.07	-0.08		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
1533	la17		419	532	1.4e-18	0.48	0.12		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
1533	la17		9	134	4.2e-16	0.14	-0.12		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE

872

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1533	1elr	A	184	284	1.1e-14	0.21	-0.02		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
1533	1elr	A	414	505	4.2e-15	0.34	0.25		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
1533	1elr	A	450	540	1.4e-13	0.36	0.30		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
1533	1elr	A	45	144	5.6e-11	0.14	0.06		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
1533	1elw	A	184	300	7e-17	0.17	0.46		TPR1-DOMAIN OF HOP; CHAIN: A; B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1533	1elw	A	423	524	4.2e-15	0.55	0.58		TPR1-DOMAIN OF HOP; CHAIN: A; B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1533	1elw	A	42	154	9.8e-10	0.20	-0.17		TPR1-DOMAIN OF HOP; CHAIN: A; B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1533	1elw	A	451	558	4.2e-15	0.58	0.07		TPR1-DOMAIN OF HOP; CHAIN: A; B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1533	1elw	A	484	591	1.4e-15	0.02	-0.19		TPR1-DOMAIN OF HOP; CHAIN: A; B; HSC70-	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL



873

Table 5

SEQ ID NO:	PDB ID	CHAI NID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PEPTIDE; CHAIN: C, D;	REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1533	1elw	A	9	120	2.8e-15	0.18	-0.14		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1533	1feh	A	219	532	3.6e-11	0.14	0.58		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISOMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
1533	1feh	A	319	598	1.3e-30	-0.00	-0.15		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	PEROXISOMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
1533	1hh8	A	10	140	7e-12	0.04	-0.12		NEUTROPHIL CYTOSOL FACTOR 2; CHAIN: A;	PHAGOCYTE OXIDASE FACTOR P67PHOX, NCF-2; PHAGOCYTE OXIDASE FACTOR, SH3 DOMAIN, REPEAT, TPR REPEAT
1533	1hxi	A	451	539	2.8e-08	0.20	-0.11		PEROXISOME TARGETING SIGNAL 1 RECEPTOR PEX5; CHAIN: A;	TRANSPORT PROTEIN PEX5; ALPHA HELICAL
1533	1ihg	A	409	538	9.8e-14	0.29	-0.08		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPIASE IMMUNOPHILIN TETRATRICOPEPTIDE
1533	1ihg	A	444	519	7.2e-07	0.04	0.63		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPIASE IMMUNOPHILIN TETRATRICOPEPTIDE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1533	1hhg	A	449	572	1.4e-12	0.15	-0.14		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPIASE IMMUNOPHILIN TETRA TRICOPEPTIDE
1533	1hhg	A	483	586	5.6e-11	0.17	-0.19		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPIASE IMMUNOPHILIN TETRA TRICOPEPTIDE
1534	1b07	A	635	667	0.0084	-0.19	0.03		PROTO-ONCOGENE CRK (CRK); CHAIN: A; HIS TAG; CHAIN: B; SH3 PEPTOID INHIBITOR; CHAIN: C;	SH3 DOMAIN P38, ADAPTER MOLECULE CRK; SH3 DOMAIN, INHIBITORS, PEPTOIDS, PROTEIN-PROTEIN 2 RECOGNITION, PROLINE-RICH MOTIFS, SIGNAL TRANSDUCTION
1534	1b0x	A	692	753	3.6e-12	0.84	1.00		EPH4 RECEPTOR TYROSINE KINASE; CHAIN: A;	TRANSFERASE RECEPTOR TYROSINE KINASE, PROTEIN INTERACTION MODULE, 2 DIMERIZATION DOMAIN, TRANSFERASE
1534	1gbr	A	635	674	0.00024	-0.01	0.06		SIGNAL TRANSDUCTION PROTEIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2, N-TERMINAL, IGBR 3 SH3 DOMAIN) COMPLEXED WITH SOS-A PEPTIDE IGBR 4 (NMR, 29 STRUCTURES) IGBR 5	
1534	1gcq	B	635	670	2.4e-06	0.13	0.00		GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; CHAIN: A, B; VAV PROTO-ONCOGENE; CHAIN: C;	SIGNALING PROTEIN/SIGNALING PROTEIN SH3 DOMAIN, PROTEIN-PROTEIN COMPLEX, GRB2, VAV

875

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1534	1hsq		635	675	0.00096	-0.39	0.10		PHOSPHORIC DIESTER HYDROLASE PHOSPHOLIPASE C-GAMMA (SH3 DOMAIN) (E.C.3.1.4.11) IHSQ 3 (NMR, MINIMIZED MEAN STRUCTURE) IHSQ 4	
1534	1pwt		635	670	0.00024	-0.05	0.07		ALPHA SPECTRIN; CHAIN: NULL;	CIRCULAR PERMUTANT PWT; CIRCULAR PERMUTANT, SH3 DOMAIN, CYTOSKELETON
1534	1sem	A	635	667	7.2e-06	-0.15	0.07		SEM-5; 1SEM 3 CHAIN: A, B; 1SEM 5 10-RESIDUE PROLINE-RICH PEPTIDE FROM MSOS 1SEM 8 CHAIN: C, D 1SEM 10	SIGNAL TRANSDUCTION PROTEIN SRC-HOMOLOGY 3 (SH3) DOMAIN, PEPTIDE-BINDING PROTEIN, 1SEM 18 2 GUANINE NUCLEOTIDE EXCHANGE FACTOR 1SEM 19
1534	1sgg		692	753	1.2e-11	0.53	0.94		EPHRIN TYPE-B RECEPTOR 2; CHAIN: NULL;	TYROSINE-PROTEIN KINASE NMR, RECEPTOR OLIGOMERIZATION, EPH RECEPTORS, TYROSINE 2 PHOSPHORYLATION, SIGNAL TRANSDUCTION, TYROSINE-PROTEIN 3 KINASE
1534	1tuc		635	676	3.6e-06	-0.55	0.13		ALPHA-SPECTRIN; CHAIN: NULL;	CYTOSKELETON CAPPING PROTEIN, CALCIUM-BINDING, DUPLICATION, REPEAT, 2 SH3 DOMAIN, CYTOSKELETON
1538	1a06		1	270	1.1e-72			108.10	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
1538	1apm	E	1	270	7e-66			70.09	TRANSFERASE(PHOSPHOTRANSFERASE) \$C- /AMP\$-DEPENDENT PROTEIN KINASE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(E.C.2.7.1.37) (\$C/APK\$) 1APM 3 (CATALYTIC SUBUNIT) "ALPHA" ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139AS) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
1538	1bi8	A	10	243	1.4e-34			62.31	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
1538	1blx	A	1	251	4.2e-37			58.85	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
1538	1cm8	A	1	270	7e-41			64.67	PHOSPHORYLATED MAP KINASE P38-GAMMA; CHAIN: A, B;	TRANSFERASE STRESS-ACTIVATED PROTEIN KINASE-3, ERK6, ERK5; P38-GAMMA, GAMMA, PHOSPHORYLATION, MAP KINASE
1538	1cmk	E	1	270	7e-67			65.16	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
1538	1cjp	E	1	270	2.8e-67			74.46	TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) ICTP 3 (CATALYTIC SUBUNIT) ICTP 4	
1538	1e1x	A	1	257	1.1e-45			75.76	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: A;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, CELL DIVISION, 2 MITOSIS, INHIBITON
1538	1f3m	C	2	262	9.8e-45			75.17	SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
1538	1fgk	A	1	244	4.2e-21			50.75	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2
										PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
1538	1fot	A	1	270	1.4e-63			87.56	CAMP-DEPENDENT PROTEIN KINASE TYPE 1; CHAIN: A;	TRANSFERASE TPK1 DELTA; CAMP-DEPENDENT PROTEIN KINASE, OPEN CONFORMATION, PROTEIN 2 KINASE
1538	1hcl		1	257	2.8e-46			66.51	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1538	1how	A	1	270	8.4e-29			57.03	SERINE/THREONINE-PROTEIN KINASE YMR216G; CHAIN: A;	CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
1538	1ia8	A	1	253	4.2e-45			82.74	CHK1 CHECKPOINT KINASE; CHAIN: A;	TRANSFERASE KINASE
1538	1iep	A	1	253	2.8e-23			60.07	PROTO-ONCOGENE TYROSINE-PROTEIN KINASE ABL; CHAIN: A, B;	TRANSFERASE P150, C-ABL; KINASE, KINASE INHIBITOR, STI-571, ACTIVATION LOOP
1538	1kob	A	1	267	7e-65			75.19	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
1538	1p38		1	270	1.4e-40			53.04	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
1538	1phk		1	173	1.4e-59			55.11	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
1538	1phk		1	246	9.8e-66			107.34	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
1538	1pme		1	270	1.4e-40			81.42	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1538	1qpc	A	1	244	1.3e-21			57.16	LCK KINASE; CHAIN: A;	KINASE, TRANSFERASE
1538	1tki	A	5	269	9.8e-57			79.61	TTTN; CHAIN: A, B;	TRANSFERASE ALPHA BETA FOLD
1538	3erk		1	266	1.4e-40			70.45	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
1540	1a1h	A	725	752	0.0024	-0.24	0.86		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE 2 ERK2
1540	1sp2		725	752	0.00024	0.25	0.71		SP1F2; CHAIN: NULL;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1540	1tf6	A	725	764	0.006	0.14	0.17		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	ZINC FINGER TRANSCRIPTION FACTOR SP1; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1
1540	1ubd	C	725	752	0.0084	-0.07	0.96		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1540	2gli	A	725	752	0.0084	0.19	0.59		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	REGULATION/DNA COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1542	lapo		911	952	2.4e-15	0.41	0.75		COAGULATION FACTOR EGF-LIKE MODULE OF BLOOD COAGULATION FACTOR X (N-TERMINAL, LAPO 3 APO FORM) (NMR, 13 STRUCTURES) LAPO 4	
1542	laui	L	915	1050	1.2e-19	0.12	-0.05		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE, PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
1542	lbf9		913	948	2.4e-15	0.58	1.00		FACTOR VII; CHAIN: NULL;	BLOOD COAGULATION BLOOD COAGULATION, EGF, HYDROLASE, SERINE PROTEASE
1542	lbf9		913	1050	8.4e-27	0.16	-0.17		PHOSPHOLIPASE A2; CHAIN: NULL;	HYDROLASE HYDROLASE, PHOSPHOLIPASE A2, PLATELET AGGREGATION INHIBITOR, 2
1542	ldan	L	1018	1108	4.8e-18	0.19	0.03		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DEFRCKM) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1542	1dan	L	913	1057	1.1e-22	0.18	0.53		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
1542	1dqb	A	913	1049	1.2e-15	0.13	-0.18		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
1542	1dqb	A	998	1095	3.6e-17	0.58	0.09		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
1542	1dva	L	913	1057	6e-23	0.13	0.18		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
1542	1dx5	I	913	1044	3.6e-26	0.21	0.15		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
1542	1edm	B	913	948	9.6e-14	0.94	1.00		FACTOR IX; CHAIN: B, C;	COAGULATION FACTOR CRYSTAL STRUCTURE,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										EPIDERMAL GROWTH FACTOR, EGF, 2 CALCIUM-BINDING, EGF-LIKE DOMAIN, STRUCTURE AND FUNCTION, 3 HUMAN FACTOR IX, COAGULATION FACTOR
1542	legf		913	950	2.4e-13	0.85	0.95		GROWTH FACTOR EPIDERMAL GROWTH FACTOR (EGF) (NMR, 16 STRUCTURES) 1EGF 3	
1542	lemn		998	1066	4.8e-19	0.16	0.84		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1542	legg	A	913	1099	8.4e-25	0.06	-0.13		PROTAGLANDIN H2 SYNTHASE-1; CHAIN: A, B;	OXIDOREDUCTASE COX-1; EGF DOMAIN, NSAID BINDING, IBUPROFEN, MEMBRANE-BINDING 2 DOMAIN, CYCLOOXYGENASE, PEROXIDASE, OXIDOREDUCTASE, 3 DIOXYGENASE, PEROXIDASE
1542	lext	A	920	1108	1.2e-26	0.37	0.24		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALING PROTEIN
1542	lfe	A	913	949	7.2e-14	1.02	1.00		BLOOD COAGULATION FACTOR VII; CHAIN: A;	BLOOD CLOTTING FACTOR VII, BLOOD COAGULATION, EGF-LIKE DOMAIN, BLOOD 2 CLOTTING
1542	lfak	L	913	1057	6e-23	0.19	0.59		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END .AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
1542	1fak	L	998	1099	1.2e-16	-0.00	0.01		BLOOD COAGULATION FACTOR VIIA; CHAIN: I; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
1542	1f6b		916	952	2.4e-15	0.83	0.87		P-SELECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN EGF-LIKE DOMAIN, CELL ADHESION PROTEIN, TRANSMEMBRANE, 2 GLYCOPROTEIN
1542	1g44	B	910	1057	3.6e-16	-0.00	-0.18		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1542	1hj7	A	1002	1080	2.4e-19	0.34	0.70		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
1542	1hj7	A	915	1045	4.8e-20	0.24	0.10		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
1542	1igr	A	913	1082	3.6e-18	0.03	-0.17		INSULIN-LIKE GROWTH FACTOR RECEPTOR I;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1542	1k1o		920	1096	2.4e-20	0.36	0.13		CHAIN: A; LAMININ; CHAIN: NULL;	FAMILY GLYCOPROTEIN GLYCOPROTEIN
1542	1pfx	L	907	1065	3.6e-18	0.16	-0.03		FACTOR IXA; CHAIN: C; L;; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
1542	1pfx	L	971	1112	4.8e-16	0.07	0.12		FACTOR IXA; CHAIN: C; L;; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
1542	1qfk	L	915	1057	6e-22	0.05	-0.17		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRUPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
1542	1tpg		907	949	6e-13	1.38	0.89		T-PLASMINOGEN ACTIVATOR FI-G; ITPG 7 CHAIN: NULL; ITPG 8	PLASMINOGEN ACTIVATION
1542	1vap	A	913	1050	6e-26	0.16	-0.18		PHOSPHOLIPASE A2; CHAIN: A, B;	LIPID DEGRADATION PHOSPHOLIPASE A2, LIPID DEGRADATION, HYDROLASE
1542	1whe		913	949	2.4e-13	0.51	0.72		COAGULATION FACTOR X; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN, HYDROLASE, SERINE PROTEASE, PLASMA, BLOOD 2

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1542	1xka	L	915	1065	4.8e-21	0.08	-0.08		BLOOD COAGULATION FACTOR XA; CHAIN: L <sub>1</sub> C <sub>1</sub>	COAGULATION FACTOR BLOOD COAGULATION FACTOR STUART FACTOR, BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
1543	1b0w	A	1	95	2.8e-57			114.58	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C <sub>1</sub>	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
1543	1b0w	A	23	130	1.1e-66			129.17	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C <sub>1</sub>	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
1543	1b2w	L	23	176	2.8e-68			110.03	ANTIBODY (LIGHT CHAIN); CHAIN: L <sub>1</sub> ANTIBODY (HEAVY CHAIN); CHAIN: H <sub>1</sub>	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
1543	1b6d	A	23	145	9.8e-72	0.83	1.00		IMMUNOGLOBULIN; CHAIN: A, B <sub>1</sub>	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
1543	1b1l	J	23	145	2.8e-73	0.78	1.00		FAB FRAGMENT; CHAIN: L <sub>1</sub> H <sub>1</sub> J <sub>1</sub> K <sub>1</sub> VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W <sub>1</sub>	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
1543	1bvk	A	1	95	1.4e-56			116.50	HULYS11; CHAIN: A, B, D, E, LYSOZYME; CHAIN: C, F <sub>1</sub>	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE, HUMANIZED ANTIBODY, ANTIBODY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
1543	1bvk	A	23	130	5.6e-67			131.09	HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
1543	1bw w	A	1	95	2.8e-59			118.29	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
1543	1bw w	A	21	129	1.4e-66			130.32	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
1543	1ce1	L	23	145	2.8e-70	0.66	1.00		CAMPATH-1H:LIGHT CHAIN; CHAIN: L; CAMPATH-1H:HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
1543	1ce1	L	23	176	2.8e-70			112.27	CAMPATH-1H:LIGHT CHAIN; CHAIN: L; CAMPATH-1H:HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
1543	1dee	A	23	145	1.4e-74	0.73	1.00		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1543	1dft	L	23	145	2.8e-69	0.78	1.00		CHAIN: G, H; IMMUNOGLOBULIN 3D6 FAB 1DFB 3	FAB VH3 3 SPECIFICITY
1543	1dql	L	1	95	7e-58			114.49	IGM MEZ IMMUNOGLOBULIN; CHAIN: L; IGM MEZ IMMUNOGLOBULIN; CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD, ANTIBODY, IGM, FV
1543	1dql	L	23	129	4.2e-65			128.96	IGM MEZ IMMUNOGLOBULIN; CHAIN: L; IGM MEZ IMMUNOGLOBULIN; CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD, ANTIBODY, IGM, FV
1543	1fgv	L	1	93	4.2e-60			118.13	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4	
1543	1fgv	L	23	129	4.2e-69	0.88	1.00		IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4	
1543	1fgv	L	23	130	4.2e-69			131.11	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4	
1543	1fvc	A	1	95	2.8e-58			119.03	IMMUNOGLOBULIN FV FRAGMENT OF HUMANIZED ANTIBODY	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	FMF score	SEQFOL D score	Compound	PDB annotation
									4D5, VERSION 8 1FVC 3	
1543	1fvc	A	23	131	2.8e-68			134.67	IMMUNOGLOBULIN FV FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 8 1FVC 3	
1543	1fvd	A	23	145	4.2e-71	0.73	1.00		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
1543	1fvd	A	23	176	4.2e-71			112.07	IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
1543	1igm	L	1	93	1.3e-58			113.94	IMMUNOGLOBULIN (IG-M) FV FRAGMENT 1IGM 3	
1543	1igm	L	23	137	2.8e-69	0.72	1.00		IMMUNOGLOBULIN (IG-M) FV FRAGMENT 1IGM 3	
1543	1igm	L	23	137	2.8e-69			132.78	IMMUNOGLOBULIN (IG-M) FV FRAGMENT 1IGM 3	
1543	1igm	L	23	131	9.8e-59			122.77	N9 NEURAMINIDASE; INMB 4 CHAIN: N; INMB 5 FAB NC10; INMB 9 CHAIN: L, H; INMB 10	COMPLEX (HYDROLASE/IMMUNOGLOBULIN)
1543	1vge	L	23	145	2.8e-69	0.70	1.00		TR1.9 FAB; CHAIN: L, H;	IMMUNOGLOBULIN TR1.9, ANTI-THYROID PEROXIDASE, AUTOANTIBODY, 2
1543	1wtl	A	1	93	5.6e-57			117.28	IMMUNOGLOBULIN	IMMUNOGLOBULIN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	FME score	SEQFOL D score	Compound	PDB annotation
									WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL.3 (BENCE-JONES PROTEIN) 1WTL.4	
1543	1wtl	A	23	130	2.8e-66			131.61	IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL.3 (BENCE-JONES PROTEIN) 1WTL.4	
									IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 1H52I (HUH52-OZ FAB) 2FGW 4	
1543	2fgw	L	23	145	1.3e-73	0.80	1.00		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN. ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
1545	1cun	A	621	803	3.6e-06	-0.08	0.51		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN
1545	1hci	A	621	803	2.4e-07	-0.18	0.09		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
1546	1a4y	A	134	363	1.2e-25	0.10	1.00		RIBONUCLEASE	COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INHIBITOR: CHAIN: A, D; ANGIOGENIN: CHAIN: B, E;	(INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPIOTOPE MAPPING, LEUCINE- RICH 3 REPEATS
1546	1a9n	A	113	213	2.8e-09	0.45	0.93		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1546	1a9n	A	122	237	3.6e-17	0.13	0.66		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1546	1a9n	A	136	265	4.8e-20	0.54	0.96		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1546	1a9n	A	158	259	1.3e-08	0.12	0.47		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1546	1a9n	A	206	343	1.1e-17	0.69	0.48		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1546	1a9n	A	255	361	4.8e-17	0.22	0.33		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1546	1a9n	C	113	213	2.8e-09	0.17	0.89		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1546	1a9n	C	136	265	2.4e-19	0.74	0.83		CHAIN: B, D; U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	SNRNP,RIBONUCLEOPROTEIN COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
1546	1a9n	C	158	259	1.3e-08	0.09	0.57		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
1546	1a9n	C	206	343	8.4e-18	0.66	0.58		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
1546	1a9n	C	255	383	6e-17	0.43	0.87		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
1546	1d0b	A	128	329	2.4e-27	0.44	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1546	1d0b	A	231	406	8.4e-21	0.20	0.63		INTERNALIN B; CHAIN: A;	CELL ADHESION LUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1546	1d0b	A	269	418	2.8e-19	0.41	0.99		INTERNALIN B; CHAIN: A;	CELL ADHESION LUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1546	1d0b	A	70	207	8.4e-20	0.39	0.98		INTERNALIN B; CHAIN: A;	CELL ADHESION LUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1546	1d0b	A	86	256	7e-24	0.52	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1546	1dee	A	158	261	1.4e-13	0.60	1.00		RAB GERANYLGERANYLTR	TRANSFERASE CRYSTAL STRUCTURE, RAB

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									NSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1546	1dce	A	227	330	1.3e-10	0.42	0.84		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1546	1dce	A	95	212	2.8e-11	0.34	0.27		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1546	1ds9	A	129	289	4.8e-17	0.18	0.82		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1546	1ds9	A	143	300	1.1e-15	-0.40	0.25		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1546	1ds9	A	210	334	2.4e-12	0.24	0.66		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, FLAGELLA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1546	1ds9	A	249	354	3.6e-16	-0.71	0.33		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1546	1ds9	A	267	383	1.3e-12	-0.23	0.24		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1546	1fo1	A	194	261	1.3e-08	-0.07	0.40		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1546	1fo1	B	194	261	1.3e-08	-0.17	0.13		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1546	1fs2	A	134	325	1.4e-11	0.22	0.36		SKP2; CHAIN: A, C, SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1546	1ft8	A	194	261	1.3e-08	0.11	0.23		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH-REPEAT 2 (LRR) DOMAINS
1546	1yrg	A	117	330	2.4e-21	-0.04	0.00		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP,

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1546	2bnh		136	363	1.2e-25	0.20	0.96		RIBONUCLEASE INHIBITOR, CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1546	2bnh		68	362	6e-25	0.11	0.94		RIBONUCLEASE INHIBITOR, CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1550	1b8q	A	69	142	5.6e-12	0.18	-0.09		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE, CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
1550	1be9	A	69	142	1.3e-20	0.24	-0.08		PSD-95; CHAIN: A; CRIP1; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
1550	1kwa	A	73	142	9.8e-13	0.00	-0.19		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN, SYNDECAN, RECEPTOR CLUSTERING, KINASE
1550	1pdt		71	142	8.4e-20	0.18	-0.09		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
1550	1gev	A	69	142	2.8e-20	0.13	-0.02		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER

895

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1550	1qlc	A	72	142	7e-18	0.01	-0.09		NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
1551	1auo	A	358	597	2.4e-30	-0.24	0.06		CARBOXYLESTERASE; CHAIN: A, B;	HYDROLASE HYDROLASE
1551	1c1g	A	74	357	2.8e-41			97.87	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
1551	1cle	A	294	595	2.8e-55	-0.17	0.00		CHOLESTEROL ESTERASE; 1CLE 4 CHAIN: A, B; 1CLE 5	LIPASE ESTERASE, SUBSTRATE/PRODUCT-BOUND 1CLE 9
1551	1dim		340	597	2.4e-23	-0.25	0.25		DIENELACTONE HYDROLASE; CHAIN: NULL;	HYDROLYTIC ENZYME DLH; DIENELACTONE HYDROLASE, AROMATIC HYDROCARBON CATABOLISM, 2 SERINE ESTERASE, CARBOXYMETHYLENEBUTENOLIDE, 3 HYDROLYTIC ENZYME
1551	1dgz	A	355	597	3.6e-30	0.05	0.30		ANTIGEN 85-C; CHAIN: A, B;	IMMUNE SYSTEM 85C ANTIGEN, 85C, MYCOBACTERIUM TUBERCULOSIS, FIBRONECTIN
1551	1ea5	A	336	597	2.4e-34	-0.06	0.11		ACETYLCHOLINESTERASE; CHAIN: A;	CHOLINESTERASE SERINE HYDROLASE, NEUROTRANSMITTER CLEAVAGE, CATALYTIC TRIAD, ALPHA/BETA HYDROLASE
1551	1evq	A	352	597	2.8e-35	0.00	0.75		SERINE HYDROLASE; CHAIN: A;	HYDROLASE ALPHA/BETA HYDROLASE FOLD
1551	1f0n	A	334	597	1.2e-32	0.04	0.64		ANTIGEN 85B; CHAIN: A;	TRANSFERASE MYCOLYL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										TRANSFERASE; 30KDA SECRETORY PROTEIN; ANTIGEN 85B
1551	1f6w	A	274	580	4.2e-64	-0.02	0.16		BILE SALT ACTIVATED LIPASE; CHAIN: A;	HYDROLASE BILE SALT ACTIVATED LIPASE, ESTERASE, CATALYTIC DOMAIN
1551	1ff2	A	349	597	4.8e-32	-0.09	0.19		ACYL PROTEIN THIOESTERASE 1; CHAIN: A, B;	HYDROLASE ALPHA/BETA HYDROLASE, SERINE HYDROLASE, SAD, ANOMALOUS 2 DIFFRACTION
1551	1jkm	A	301	575	2.8e-26	-0.02	0.80		BREFELDIN A ESTERASE; CHAIN: A, B;	SERINE HYDROLASE SERINE HYDROLASE, DEGRADATION OF BREFELDIN A, ALPHA/BETA 2 HYDROLASE FAMILY
1551	1lpp		294	595	1.4e-54	-0.30	0.00		HYDROLASE LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL LIPASE) COMPLEXED WITH 1LPP 3 HEXADECANESULFONATE 1LPP 4 1LPP 71	
1551	1maa	A	336	592	2.4e-34	-0.06	0.25		ACETYLCHOLINESTERASE; CHAIN: A, B, C, D;	HYDROLASE MACHE; HYDROLASE, SERINE ESTERASE, ACETYLCHOLINESTERASE, TETRAMER, 2 HYDROLASE FOLD, GLYCOSYLATED PROTEIN
1551	1qfm	A	1	597	5.6e-81			90.26	PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
1551	1qfm	A	3	597	5.6e-81	-0.04	0.89		PROLYL OLIGOPEPTIDASE;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A;	CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
1551	1thg		284	594	2.8e-59	0.00	0.59		HYDROLASE(CARBOXY LIC ESTERASE) LIPASE (E.C.3.1.1.3) TRIACYLGLYCEROL HYDROLASE 1THG 3	
1552	1d7k	A	18	220	1.4e-59	0.22	0.77		HUMAN ORNITHINE DECARBOXYLASE; CHAIN: A, B;	LYASE ALPHA-BETA BARREL, PYRIDOXAL 5'-PHOSPHATE, SHEET-DOMAIN, 2 DECARBOXYLATION, ORNITHINE
1552	1f3t	A	18	203	1.4e-51	-0.09	0.96		ORNITHINE DECARBOXYLASE; CHAIN: A, B, C, D;	LYASE ODC, BETA-ALPHA-BARREL, MODIFIED GREEK KEY BETA-SHEET
1552	2tod	A	18	203	2.8e-51	-0.10	0.96		ORNITHINE DECARBOXYLASE; CHAIN: A, B, C, D;	LYASE POLYAMINE METABOLISM, PYRIDOXAL 5'-PHOSPHATE, ALPHA-BETA 2 BARREL, LYASE
1552	7odc	A	18	210	2.8e-60	0.08	0.81		ORNITHINE DECARBOXYLASE; CHAIN: A;	LYASE ODC, MODC, MODC; PYRIDOXAL-5'-PHOSPHATE, PLP, GROUP IV DECARBOXYLASE, 2 POLYAMINES, PARASITIC, CHEMOTHERAPY TARGET, ORNITHINE, 3 PUTRESCINE, A/B-BARREL, OBLIGATE, LYASE
1556	1ad0	A	382	437	0.00012	-0.61	0.30		FAB FRAGMENT, ANTIBODY A5B7; CHAIN: A, B, C, D;	IMMUNOGLOBULIN FAB FRAGMENT
1556	1b4j	L	382	437	0.00012	-0.54	0.10		ANTIBODY; CHAIN: L, H;	ANTIBODY ENGINEERING ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODIES, 2 FAB, X-RAY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1556	1dee	A	382	437	0.00012	-0.42	0.55		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	STRUCTURES, GAMMA-INTERFERON
1556	1igt	A	382	437	8.4e-05	-0.61	0.22		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1556	3fct	A	382	437	9.6e-05	-0.63	0.41		METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: A, C; METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: B, D;	IMMUNE SYSTEM METAL CHELATASE, CATALYTIC ANTIBODY, FAB FRAGMENT, IMMUNE 2 SYSTEM
1557	1rth	A	206	320	0.0024	-0.42	0.39		HIV-1 REVERSE TRANSCRIPTASE; 1RTH 4 CHAIN: A, B; 1RTH 5	NUCLEOTIDYLTRANSFERASE HIV-1 RT; 1RTH 6 HIV-1 REVERSE TRANSCRIPTASE 1RTH 15
1557	1rth	B	206	320	0.0012	-0.32	0.51		HIV-1 REVERSE TRANSCRIPTASE; 1RTH 4 CHAIN: A, B; 1RTH 5	NUCLEOTIDYLTRANSFERASE HIV-1 RT; 1RTH 6 HIV-1 REVERSE TRANSCRIPTASE 1RTH 15
1557	1vrt	A	206	320	0.0024	-0.35	0.53		HIV-1 REVERSE TRANSCRIPTASE; 1VRT 4 CHAIN: A, B; 1VRT 5	NUCLEOTIDYLTRANSFERASE HIV-1 RT; 1VRT 6 HIV-1 REVERSE TRANSCRIPTASE 1VRT 15
1557	1vrt	B	206	320	0.0012	-0.54	0.36		HIV-1 REVERSE TRANSCRIPTASE; 1VRT 4 CHAIN: A, B; 1VRT 5	NUCLEOTIDYLTRANSFERASE HIV-1 RT; 1VRT 6 HIV-1 REVERSE TRANSCRIPTASE 1VRT 15
1559	1byg	A	149	397	1.4e-30	0.02	0.24		C-TERMINAL SRC KINASE; CHAIN: A;	TRANSFERASE CSK; PROTEIN KINASE, C-TERMINAL SRC KINASE, PHOSPHORYLATION, 2
1559	1cki	A	149	438	2.4e-20	0.10	0.05		CASEIN KINASE 1	STAUROSPORINE, TRANSFERASE PHOSPHOTRANSFERASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DELTA; ICKI 6 CHAIN: A, B; ICKI 7	PROTEIN KINASE ICKI 18
1559	1f3m	C	130	464	1.4e-81	-0.14	0.12		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
1559	1f6k	A	137	384	2.8e-32	0.26	-0.01		FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2
										PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
1559	1f6k	B	137	384	7e-36	0.09	0.74		FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2
										PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
1559	1f6k		146	367	1.4e-32	0.22	0.04		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2
										PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
1559	1f6k		375	459	1.1e-08	-0.16	0.03		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2
										PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
1559	1h0w	A	148	231	2.4e-11	0.32	-0.01		SERINE/THREONINE-	TRANSFERASE KINASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEIN KINASE YMK216C; CHAIN: A;	
1559	1i44	A	149	231	1.2e-09	-0.26	0.47		INSULIN RECEPTOR; CHAIN: A;	TRANSFERASE IR; PROTEIN TYROSINE KINASE, PHOSPHOTRANSFERASE
1559	1ia8	A	149	462	8.4e-54	0.04	0.29		CHK1 CHECKPOINT KINASE; CHAIN: A;	TRANSFERASE PROTEIN KINASE
1559	1iep	A	156	367	5.6e-33	0.01	0.24		PROTO-ONCOGENE TYROSINE-PROTEIN KINASE ABL; CHAIN: A, B;	TRANSFERASE P150, C-ABL; KINASE, KINASE INHIBITOR, STI-571, ACTIVATION LOOP
1559	1kob	A	149	464	1.4e-53	-0.02	0.40		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
1559	1p38		142	231	1.2e-12	0.14	0.36		MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
1559	1qcf	A	146	433	2.8e-34	0.08	-0.06		HAEMATOPHOETIC CELL KINASE (HCK); CHAIN: A;	TYROSINE KINASE TYROSINE KINASE-INHIBITOR COMPLEX, DOWN-REGULATED KINASE, 2 ORDERED ACTIVATION LOOP
1559	1qpc	A	147	374	2.8e-32	0.22	0.43		LCK KINASE; CHAIN: A;	TRANSFERASE ALPHA BETA FOLD
1559	1vr2	A	140	367	9.8e-34	0.05	0.15		VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR CHAIN: A;	TRANSFERASE KDR; TYROSINE KINASE
1559	1vr2	A	365	460	2.8e-07	-0.22	0.13		VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR CHAIN: A;	TRANSFERASE KDR; TYROSINE KINASE
1560	1atl	A	9	214	1.4e-70			94.37	ATROLYSIN C; LATL 4 CHAIN: A, B, C, D, LATL	METALLOENDOPEPTIDASE HEMORRHAGIC TOXIN C, FORM

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1560	1bkc	A	5	213	1.4e-24			52.22	5	D-1ATL 6
										ZN-ENDOPEPTIDASE TACE; ZN-ENDOPEPTIDASE, HYDROLASE, TNF-ALPHA
1560	1bud	A	9	212	4.2e-67			77.02	ACUTOLYSIN A; CHAIN: A;	TOXIN HEMORRHAGIN I, IAAH-I; METALLOPROTEINASE, SNAKE VENOM, MMP, TOXIN
1560	1cii		506	1002	2.4e-12	-0.41	0.13		COLICIN IA; CHAIN: NULL;	TRANSMEMBRANE PROTEIN COLICIN, BACTERIOCIN, ION CHANNEL FORMATION, TRANSMEMBRANE 2 PROTEIN
1560	1cun	A	636	896	0.00072	-0.03	0.05		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
1560	1dd5	A	36	215	0.007			51.70	RIBOSOME RECYCLING FACTOR; CHAIN: A;	RIBOSOME THREE-HELIX BUNDLE, BETA-ALPHA-BETA SANDWICH, RIBOSOME
1560	1lci	A	514	1000	1.2e-12	-0.42	0.18		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN
1560	1lci	A	514	983	1.2e-12			140.65	ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1560	1hci	A	638	1060	2.4e-07	-0.12	0.68		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN-BINDING PROTEIN
1560	1htp		6	133	1.1e-52			114.21	OXIDOREDUCTASES(AC TING ON CH-NH2 DONOR) H-PROTEIN (E.C.1.4.4.2) COMPLEXED WITH LIPOLIC ACID CHARGED 1HTP 3 IN METHYLAMINE 1HTP 4 1HTP 77	
1560	1hzp	A	3	356	8.4e-56			72.15	3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE III; CHAIN: A, B;	TRANSFERASE MT-FABH; FATTY ACID BIOSYNTHESIS, MYOBACTERIUM TUBERCULOSIS, 2 STRUCTURAL BASIS FOR SUBSTRATE SPECIFICITY
1560	1i84		S114	1008	1.2e-13			158.73	SMOOTH MUSCLE MYOSIN HEAVY CHAIN; CHAIN: S, V; SMOOTH MUSCLE MYOSIN ESSENTIAL LIGHT CHAIN; CHAIN: T, W; SMOOTH MUSCLE MYOSIN REGULATORY LIGHT CHAIN; CHAIN: U, Z;	CONTRACTILE PROTEIN MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSCLE; ELC, MYOSIN ALKALI LIGHT CHAIN; RLC, MYOSIN REGULATORY LIGHT CHAIN 2, SMOOTH MUSCLE PROTEIN, SMOOTH MUSCLE, MYOSIN SUBFRAGMENT 2, HEAVY 2 MEROMYOSIN, ESSENTIAL LIGHT CHAIN, REGULATORY LIGHT CHAIN, 3 MOTOR PROTEIN, COILED-COIL
1560	1iag		5	214	1.4e-67			82.71	METALLOPROTEASE ADAMALYSIN II	

903

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1560	1qua	A	5	212	4.2e-65			84.37	(PROTEINASE II) (E.C.3.4.24.46) IIAG 3 ACUTOLYSIN-C; CHAIN: A;	TOXIN HEMORRHAGIN III METALLOPROTEASE, HEMORRHAGIC TOXIN, SNAKE VENOM PROTEINASE, 2 CRYSTAL STRUCTURE, AGKISTRODON ACUTUS
1560	1quu	A	525	796	0.0024	-0.27	0.05		HUMAN SKELETAL MUSCLE ALPHA- ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE- HELIX COILED COIL, CONTRACTILE PROTEIN
1560	1ses	A	872	944	4.2e-05	0.06	-0.03		LIGASE(SYNTHETASE) SERYL-TRNA SYNTHETASE (E.C.6.1.1.11) (SERINE- TRNA LIGASE) 1SES 3 COMPLEXED WITH SERYL- HYDROXYAMATE-AMP 1SES 4	
1560	1sig		514	778	6e-06	-0.29	0.05		RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
1568	1hbp		60	187	4.8e-35			113.26	OXIDOREDUCTASES(AC TING ON CH-NH2 DONOR) H-PROTEIN (E.C.1.4.4.2) COMPLEXED WITH LIPIC ACID CHARGED 1HTP 3 IN METHYLAMINE 1HTP 4 1HTP 77	
1568	1hbp		65	183	4.8e-35	1.41	1.00		OXIDOREDUCTASES(AC TING ON CH-NH2 DONOR) H-PROTEIN	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1568	1htp		65	185	7e-27	1.38	1.00		(E.C.1.4.4.2) COMPLEXED WITH LIPOIC ACID CHARGED 1HTP 3 IN METHYLAMINE 1HTP 4 1HTP 77	
									OXIDOREDUCTASES(AC TING ON CH-NH2 DONOR) H-PROTEIN (E.C.1.4.4.2) COMPLEXED WITH LIPOIC ACID CHARGED 1HTP 3 IN METHYLAMINE 1HTP 4 1HTP 77	
1570	1a88	A	176	385	0.0096	-0.10	0.29		CHLOROPEROXIDASE L; CHAIN: A, B, C;	HALOPEROXIDASE BROMOPEROXIDASE L, HALOPEROXIDASE L; HALOPEROXIDASE, HALOPEROXIDASE, OXIDOREDUCTASE
1570	1a8s		177	385	0.006	0.10	0.25		CHLOROPEROXIDASE F; CHAIN: NULL;	HALOPEROXIDASE HALOPEROXIDASE F; HALOPEROXIDASE, OXIDOREDUCTASE, PROPIONATE COMPLEX
1570	1dim		188	431	0.006	0.11	0.10		DIENELACTONE HYDROLASE; CHAIN: NULL;	HYDROLYTIC ENZYME DLH; DIENELACTONE HYDROLASE, AROMATIC HYDROCARBON CATABOLISM, 2 SERINE ESTERASE, CARBOXYMETHYLENEBUTENOLIDE, 3 HYDROLYTIC ENZYME
1572	2gli	A	586	647	1.4e-13	0.30	0.23		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1572	2gli	A	626	652	4.2e-10	0.29	0.90		ZINC FINGER PROTEIN	COMPLEX (DNA-BINDING



905

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GLI1; CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1575	1qfh	A	536	641	0.00096	-0.28	0.15		GELATION FACTOR; CHAIN: A, B;	ACTIN BINDING PROTEIN ACTIN BINDING PROTEIN 120; ACTIN BINDING PROTEIN, IMMUNOGLOBULIN, GELATION FACTOR, ABP-2 120
1577	1hs6	A	34	614	4.8e-77			110.77	LEUKOTRIENE A-4 HYDROLASE; CHAIN: A;	HYDROLASE PROTEIN-INHIBITOR COMPLEX, ALPHA-BETA PROTEIN
1577	1hs6	A	41	594	4.8e-77	0.24	0.47		LEUKOTRIENE A-4 HYDROLASE; CHAIN: A;	HYDROLASE PROTEIN-INHIBITOR COMPLEX, ALPHA-BETA PROTEIN
1577	1hs6	A	47	505	1.3e-78	0.22	0.69		LEUKOTRIENE A-4 HYDROLASE; CHAIN: A;	HYDROLASE PROTEIN-INHIBITOR COMPLEX, ALPHA-BETA PROTEIN
1579	1bf7f	A	821	939	1.3e-15	0.26	0.95		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(p*Gp*Up*Up*Gp*Up*Up*Up*Up*Up*Up*U)-CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
1579	1cvj	A	823	934	1.4e-15	0.47	0.65		POLYDENYLATE BINDING PROTEIN I; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*Ap*Ap*Ap*Ap*Ap*A p*Ap*Ap*Ap*Ap*A)-3); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN I, PABP I; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
1579	1cvj	B	721	878	1.4e-30	0.01	-0.17		POLYDENYLATE BINDING PROTEIN I; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-	GENE REGULATION/RNA POLY(A) BINDING PROTEIN I, PABP I; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									R(*AP*AP*AP*AP*AP*A P*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	
1579	1cvj	F	823	938	9.8e-16	0.30	0.39		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'- R(*AP*AP*AP*AP*AP*A P*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
1579	1cvj	H	823	938	9.8e-16	0.26	0.72		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'- R(*AP*AP*AP*AP*AP*A P*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
1579	1fxl	A	821	940	5.6e-19	0.17	0.99		PARANEOPLASTIC ENCEPHALOMYEELITIS ANTIGEN HUD; CHAIN: A; 5'- R(P*UP*UP*UP*UP*AP*UP*UP*UP*U)-3'; CHAIN: B;	TRANSCRIPTION/RNA HUD, HU- ANTIGEN D, PROTEIN-RNA COMPLEX, HUD, AU-RICH ELEMENT
1579	1ha1		721	894	2.8e-42	0.09	-0.12		HNRRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
1579	1ha1		821	939	2.8e-25	0.14	0.89		HNRRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
1579	1hdc0	A	822	894	8.4e-17	0.31	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
1579	2mss	A	822	894	2.8e-15	0.34	0.94		MUSASHI1; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
1579	2up1	A	721	894	2.8e-44	0.09	-0.11		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UPI; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
1579	2up1	A	821	939	1.1e-24	0.27	0.89		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UPI; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
1579	3sxl	A	821	939	9.8e-16	0.16	0.93		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
1580	1a7c	A	267	544	2.8e-62	-0.24	0.12		PLASMINOGEN	COMPLEX (PROTEASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1580	1a7c	A	34	337	7e-72	0.04	1.00		ACTIVATOR INHIBITOR TYPE 1; CHAIN: A; PENTAPEPTIDE; CHAIN: B, C	INHIBITOR/PEPTIDE) PAI-1; SERINE PROTEASE INHIBITOR, PAI-1, CARBOHYDRATE, INHIBITOR 2 COMPLEX, PEPTIDE
1580	1a7c	A	34	337	7e-72	0.04	1.00		PLASMINOGEN ACTIVATOR INHIBITOR TYPE 1; CHAIN: A; PENTAPEPTIDE; CHAIN: B, C	COMPLEX (PROTEASE INHIBITOR/PEPTIDE) PAI-1; SERINE PROTEASE INHIBITOR, PAI-1, CARBOHYDRATE, INHIBITOR 2 COMPLEX, PEPTIDE
1580	1a7c	A	8	225	1.4e-62			59.79	PLASMINOGEN ACTIVATOR INHIBITOR TYPE 1; CHAIN: A; PENTAPEPTIDE; CHAIN: B, C	COMPLEX (PROTEASE INHIBITOR/PEPTIDE) PAI-1; SERINE PROTEASE INHIBITOR, PAI-1, CARBOHYDRATE, INHIBITOR 2 COMPLEX, PEPTIDE
1580	1as4	A	39	374	1.1e-80			94.64	ANTICHYMOTRYPSIN; CHAIN: A, B;	SERPIN ACT; SERPIN, SERINE PROTEASE INHIBITOR, ANTICHYMOTRYPSIN
1580	1as4	A	40	309	1.1e-80	0.04	1.00		ANTICHYMOTRYPSIN; CHAIN: A, B;	SERPIN ACT, SERPIN, SERINE PROTEASE INHIBITOR, ANTICHYMOTRYPSIN
1580	1ath	A	31	405	2.8e-66			132.53	HUMAN ANTIHROMBIN-III ANTIHROMBIN III IATH 3	
1580	1ath	B	30	388	4.2e-67			119.29	HUMAN ANTIHROMBIN-III ANTIHROMBIN III IATH 3	
1580	1ath	B	38	309	4.2e-67	0.07	1.00		HUMAN ANTIHROMBIN-III ANTIHROMBIN III IATH 3	
1580	1by7	A	38	402	7e-66			104.95	PLASMINOGEN ACTIVATOR INHIBITOR-	PROTEIN BINDING PAI-2; SERPIN, PROTEIN BINDING

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1580	1db2	A	267	544	1.4e-68	0.01	0.59		2; CHAIN: A; PLASMINOGEN ACTIVATOR INHIBITOR-1; CHAIN: A, B;	HYDROLASE INHIBITOR NATIVE SERPIN, HYDROLASE INHIBITOR
1580	1db2	A	35	335	4.2e-71	0.02	0.96		PLASMINOGEN ACTIVATOR INHIBITOR-1; CHAIN: A, B;	HYDROLASE INHIBITOR NATIVE SERPIN, HYDROLASE INHIBITOR
1580	1db2	A	9	225	2.8e-61			50.63	PLASMINOGEN ACTIVATOR INHIBITOR-1; CHAIN: A, B;	HYDROLASE INHIBITOR NATIVE SERPIN, HYDROLASE INHIBITOR
1580	1d2g	I	1	405	1.4e-67			136.83	ANTITHROMBIN-III; CHAIN: I; ANTITHROMBIN-III; CHAIN: I;	SERPIN SERPIN
1580	1d2g	I	6	306	1.4e-67	-0.03	0.92		ANTITHROMBIN-III; CHAIN: I; ANTITHROMBIN-III; CHAIN: I;	SERPIN SERPIN
1580	1ezx	A	39	307	2.8e-80	-0.06	0.90		ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1- ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
1580	1ezx	A	39	375	2.8e-80			110.88	ALPHA-1-ANTITRYPSIN; CHAIN: A; ALPHA-1- ANTITRYPSIN; CHAIN: B; TRYPSIN; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR PROTEASE-INHIBITOR COMPLEX, SERPIN, ALPHA-1-ANTITRYPSIN, 2 TRYPSIN
1580	1hle	A	37	309	1.4e-77	0.25	0.94		HYDROLASE INHIBITOR(SERINE PROTEINASE) HORSE LEUKOCYTE ELASTASE INHIBITOR (FILED) 1HLE 3	
1580	1hle	A	38	378	1.4e-77			111.65	HYDROLASE INHIBITOR(SERINE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEINASE) HORSE LEUKOCYTE ELASTASE INHIBITOR (HLEI) 1HLE 3	
1580	1i99	I	267	544	9.8e-61	-0.26	0.24		ALASERPIN; CHAIN: I; TRYPSIN II, ANIONIC; CHAIN: E;	HYDROLASE/HYDROLASE INHIBITOR SERPIN I; PRETRYPSINOGEN II; MICHAELIS SERPIN-PROTEASE COMPLEX INHIBITORY TRIAD
1580	1i99	I	352	537	7.2e-31	-0.12	0.18		ALASERPIN; CHAIN: I; TRYPSIN II, ANIONIC; CHAIN: E;	HYDROLASE/HYDROLASE INHIBITOR SERPIN I; PRETRYPSINOGEN II; MICHAELIS SERPIN-PROTEASE COMPLEX INHIBITORY TRIAD
1580	1ova	A	234	544	5.6e-70	-0.24	0.51		SERPIN OVALBUMIN (EGG ALBUMIN) 1OVA 3	
1580	1ova	A	39	304	5.6e-67	0.03	1.00		SERPIN OVALBUMIN (EGG ALBUMIN) 1OVA 3	
1580	1ova	A	39	402	5.6e-67			108.90	SERPIN OVALBUMIN (EGG ALBUMIN) 1OVA 3	
1580	1qlp	A	35	390	2.8e-80			107.96	ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE; SERINE PROTEASE INHIBITOR, SERPIN, GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
1580	1qlp	A	39	307	2.8e-80	-0.02	0.89		ALPHA-1-ANTITRYPSIN; CHAIN: A;	SERINE PROTEASE INHIBITOR ALPHA-1-PROTEINASE INHIBITOR, ALPHA-1-ANTIPROTEINASE; SERINE PROTEASE INHIBITOR, SERPIN,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										GLYCOPROTEIN, SIGNAL, 2 POLYMORPHISM, EMPHYSEMA, DISEASE MUTATION, ACUTE PHASE
1580	1qmn	A	181	544	9.8e-65	-0.06	0.11		ALPHA-1-ANTICHYMOTRYPSIN; CHAIN: A;	SERPIN AACT SERPIN, SERINE PROTEINASE INHIBITOR, PARTIAL: LOOP 2 INSERTION, LOOP-SHEET POLYMERIZATION, EMPHYSEMA, DISEASE 3 MUTATION, ACUTE PHASE PROTEIN, CONFORMATIONAL DISEASE
1580	1qmn	A	40	309	1.1e-78	0.04	0.99		ALPHA-1-ANTICHYMOTRYPSIN; CHAIN: A;	SERPIN AACT SERPIN, SERINE PROTEINASE INHIBITOR, PARTIAL: LOOP 2 INSERTION, LOOP-SHEET POLYMERIZATION, EMPHYSEMA, DISEASE 3 MUTATION, ACUTE PHASE PROTEIN, CONFORMATIONAL DISEASE
1580	1qmn	A	40	404	1.1e-78			118.53	ALPHA-1-ANTICHYMOTRYPSIN; CHAIN: A;	SERPIN AACT SERPIN, SERINE PROTEINASE INHIBITOR, PARTIAL: LOOP 2 INSERTION, LOOP-SHEET POLYMERIZATION, EMPHYSEMA, DISEASE 3 MUTATION, ACUTE PHASE PROTEIN, CONFORMATIONAL DISEASE
1580	1sek		267	544	1.1e-59	-0.29	0.24		SERPIN K; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, SERPIN, PROTEASE
1580	1sek		33	388	2.8e-63			140.59	SERPIN K; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, SERPIN, PROTEASE
1582	1a4y	A	188	493	1.4e-12	-0.03	0.60		RIBONUCLEASE	COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	(INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1582	1a4y	A	54	371	2.4e-36	-0.13	0.18		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1582	1a9n	A	171	320	1.2e-21	0.51	0.42		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	A	192	298	4.2e-07	0.71	0.47		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	A	220	372	1.2e-18	0.25	0.07		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	A	247	378	8.4e-20	0.29	0.30		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	A	288	396	0.00014	-0.17	0.11		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1582	1a9n	A	51	134	2.4e-07	0.06	0.41		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	A	51	154	5.6e-07	0.11	0.31		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	A	76	218	4.8e-26	0.46	0.54		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	C	171	320	2.4e-22	0.37	0.39		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	C	192	298	4.2e-07	0.58	0.36		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	C	220	372	2.4e-18	0.26	0.05		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	C	247	378	4.8e-20	0.16	0.22		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	C	288	396	0.00014	0.15	0.46		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	C	51	154	5.6e-07	0.20	0.34		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQROL D score	Compound	PDB annotation
									CHAIN: A, C; U2 B"; CHAIN: B, D;	(NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1a9n	C	76	250	9.6e-28	0.38	0.34		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A"; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1582	1cs6	A	429	601	2.8e-14	0.30	0.29		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1582	1cvs	C	440	603	4.2e-11	0.06	-0.09		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1582	1cvs	D	430	514	4.8e-15	-0.06	0.75		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1582	1d0b	A	116	320	8.4e-27	0.52	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1582	1d0b	A	162	369	9.6e-23	0.17	0.83		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1582	1d0b	A	236	416	2.8e-19	0.17	0.99		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1582	1d0b	A	79	268	2.8e-26	0.34	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1582	1dce	A	185	273	2.8e-12	0.57	1.00		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1582	1dce	A	216	340	2.8e-13	0.37	0.96		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1582	1dce	A	288	395	1.4e-08	0.16	0.95		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1582	1dce	A	42	129	4.2e-09	-0.28	0.84		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1582	1dce	A	51	268	3.6e-25	-0.32	0.00		RAB GERANYLGERANYLTR NSFERASE ALPHA	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS

916

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	E, 2.0 Å 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1582	1ds9	A	199	362	8.4e-15	-0.13	0.48		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1582	1ds9	A	225	365	1.4e-12	-0.43	0.19		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1582	1ds9	A	239	371	2.4e-14	-0.32	0.10		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1582	1ds9	A	45	153	8.4e-11	-0.21	0.04		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1582	1ds9	A	92	274	7.2e-23	-0.16	0.15		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1582	1epf	A	432	514	3.6e-15	0.28	0.89		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM, NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1582	1ev2	E	439	586	5.6e-12	0.15	-0.05		FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1582	lev2	G	439	586	7e-12	0.08	-0.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1582	levt	C	440	586	1.4e-11	0.09	-0.11		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1582	lfhg	A	415	514	2.4e-18	-0.15	0.46		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
1582	lfna		531	593	0.006	0.22	0.06		CELL ADHESION PROTEIN FIBRONECTIN CELL-ADHESION MODULE TYPE III-10 1FN A 3	
1582	lfnf		454	595	2.8e-05	0.21	-0.01		FIBRONECTIN; 1FN F 6 CHAIN: NULL; 1FN F 7	CELL ADHESION PROTEIN RGD, EXTRACELLULAR MATRIX 1FN F 18
1582	lfo1	A	308	371	0.00012	-0.53	0.17		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1582	lfo1	A	50	147	1.4e-07	-0.19	0.24		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										(RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1582	1fo1	B	333	425	1.1e-07	-0.01	0.17		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1582	1fo1	B	50	147	1.4e-07	-0.06	0.24		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1582	1fo1	B	91	194	2.4e-13	-0.68	0.12		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1582	1fgv	A	192	368	1.4e-15	0.48	0.75		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1582	1fs2	A	110	362	4.8e-16	-0.05	0.09		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1582	1fs2	A	192	368	1.4e-15	0.40	0.89		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1582	1fs2	A	72	288	1.2e-12	-0.11	0.06		SKP2; CHAIN: A, C; SKP1;	LIGASE CYCLIN A/CDK2-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1582	1ff8	A	333	425	1.1e-07	-0.06	0.11		CHAIN: B, D; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRKS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1582	1ff8	A	50	147	1.4e-07	-0.06	0.09		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH-REPEAT 2 (LRR) DOMAINS
1582	1ff8	A	50	147	1.4e-07	-0.06	0.09		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E; TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH-REPEAT 2 (LRR) DOMAINS
1582	1ie5	A	422	514	6e-19	0.07	0.24		NEURAL CELL ADHESION MOLECULE; CHAIN: A; HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	CELL ADHESION N-CAM; INTERMEDIATE IMMUNOGLOBULIN FOLD
1582	1iil	G	422	514	1.2e-15	0.11	0.98		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREEFOIL
1582	1iil	G	439	586	5.6e-11	-0.05	0.00		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREEFOIL
1582	1mfn		504	597	0.0059	0.22	-0.06		FIBRONECTIN; CHAIN: NULL; FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1582	Inct		422	514	1.2e-17	0.06	0.19		TTTN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
1582	Iten		531	593	0.0096	0.12	0.05		CELL ADHESION PROTEIN TENASCIN (THIRD FIBRONECTIN TYPE III REPEAT) ITEN 3	
1582	Itmn		430	514	1.1e-17	0.08	0.37		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) ITNM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) ITNM 4 ITNM 58	
1582	Ivca	A	430	585	4.8e-11	0.11	0.18		HUMAN VASCULAR CELL ADHESION MOLECULE-1; IVCA 4 CHAIN: A, B; IVCA 5	CELL ADHESION PROTEIN VCAM-D1,2; IVCA 6 IMMUNOGLOBULIN SUPERFAMILY, INTEGRIN-BINDING IVCA 15
1582	Iwit		430	514	1.2e-15	0.27	-0.07		TWITCHIN 18TH IGSP MODULE; CHAIN: NULL;	MUSCLE PROTEIN IMMUNOGLOBULIN SUPERFAMILY, ISET, MUSCLE PROTEIN
1582	Iyrg	A	180	372	4.2e-11	0.27	0.87		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMISPHERAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1582	Iyrg	A	40	202	1.4e-06	-0.32	0.23		GTPASE-ACTIVATING	TRANSCRIPTION RNA1P;



921

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEIN RNAI_SCHPO; CHAIN: A, B;	RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE- ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1582	1yrg	A	72	202	6e-14	0.01	0.05		GTPASE-ACTIVATING PROTEIN RNAI_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE- ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1582	1yrg	A	87	320	2.4e-30	0.39	0.86		GTPASE-ACTIVATING PROTEIN RNAI_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE- ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1582	2bnh		188	472	4.2e-15	0.11	0.62		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1582	2bnh		40	457	1.4e-20	-0.06	0.21		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION,

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1582	2bnh		55	375	2.4e-37	0.07	0.39		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	LEUCINE-RICH REPEATS ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1582	2ncm		430	514	4.8e-17	0.26	0.98		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
1582	3ncm	A	430	514	8.4e-18	0.45	0.83		NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM; CHAIN: A;	CELL ADHESION PROTEIN NCAM MODULE 2; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, HOMOPHILIC 3 BINDING, CELL ADHESION PROTEIN
1585	1lxa		364	470	3.6e-05	1.02	0.36		UDP N-ACETYLGLUCOSAMINE O-ACYLTRANSFERASE; ILXA 5 CHAIN: NULL; ILXA 6	ACYLTRANSFERASE LPXA; ILXA 7 TRANSFERASE, ACYLTRANSFERASE, LIPID A BIOSYNTHESIS, ILXA 17 2 LIPID SYNTHESIS ILXA 18
1585	1gre	A	364	470	0.00012	0.73	0.65		CARBONIC ANHYDRASE; CHAIN: A;	LYASE BETA-HELIX
1585	1xat		364	475	0.0072	0.70	-0.09		XENOBIOTIC ACETYLTRANSFERASE; CHAIN: NULL;	ACETYLTRANSFERASE ACETYLTRANSFERASE, XENOBIOTIC, CHLORAMPHENICOL, LEFT-HANDED 2 BETA HELIX
1587	1bor		171	208	0.00028	-0.25	0.09		TRANSCRIPTION	TRANSCRIPTION REGULATION

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOLD score	Compound	PDB annotation
1591	1a17		252	393	2.8e-30	0.09	1.00		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	PROTO-ONCOGENE, NUCLEAR BODIES (PODS), LEUKEMIA, 2 TRANSCRIPTION REGULATION
1591	1b3u	A	142	751	2.4e-13			112.25	PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	HYDROLASE TETRAPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
1591	1b3u	A	193	804	2.4e-13	0.03	0.88		PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT REPEAT
1591	1c1g	A	1	249	8.4e-05			50.04	TROPOMYOSIN; CHAIN: A, B, C, D	SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT REPEAT
1591	1ee4	A	396	800	3.6e-16	0.47	1.00		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
1591	1ee4	A	635	816	1.2e-08	0.11	0.83		KARYOPHERIN ALPHA; CHAIN: A, B; MYC PROTO-ONCOGENE PROTEIN; CHAIN: C, D, E, F;	TRANSPORT PROTEIN SERINE-RICH RNA POLYMERASE I SUPPRESSOR PROTEIN; ARM REPEAT
1591	1elt	A	252	371	2.8e-20	0.13	0.99		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1591	1elr	A	256	371	1.2e-24	0.38	0.98		CHAIN: B; TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90- PEPTIDE MEEVD; CHAIN: B;	BINDING CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
1591	1elw	A	231	322	4.2e-13	-0.34	0.13		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70- PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1591	1elw	A	252	368	8.4e-24	0.33	1.00		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70- PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1591	1elw	A	252	371	1.1e-26	0.32	1.00		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70- PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1591	1elw	A	299	399	1.4e-11	-0.51	0.23		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70- PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1591	1fch	A	231	477	1.4e-31	-0.30	0.05		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISOMAL RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRAPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
1591	1g3j	C	461	802	2.4e-14	0.13	0.46		BETA-CATENIN ARMADILLO REPEAT REGION; CHAIN: A, C; TCF3-CBD (CATENIN BINDING DOMAIN); CHAIN: B, D;	TRANSCRIPTION BETA- CATENIN, TCF-3, PROTEIN- PROTEIN COMPLEX
1591	1g3j	C	476	796	8.4e-11	0.34	0.64		BETA-CATENIN	TRANSCRIPTION BETA-

925

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ARMADILLO REPEAT REGION; CHAIN: A, C; TCF3-CBD (CATENIN BINDING DOMAIN); CHAIN: B, D;	CATENIN,TCF-3, PROTEIN-PROTEIN COMPLEX
1591	1hh8	A	253	443	2.8e-15	-0.30	0.19		NEUTROPHIL CYTOSOL FACTOR 2; CHAIN: A;	PHAGOCYTE OXIDASE FACTOR P67PHOX, NCF-2; PHAGOCYTE OXIDASE FACTOR, SH3 DOMAIN, REPEAT, TPR REPEAT
1591	1i7w	A	382	757	4.8e-07	0.10	0.77		BETA-CATENIN; CHAIN: A, C; EPTHELIAL-CADHERIN; CHAIN: B, D;	CELL ADHESION E-CADHERIN; E-CADHERIN, CELL ADHESION, BETA-CATENIN, PROTEIN-PROTEIN 2 COMPLEX, EXTENDED INTERFACE, ARMADILLO REPEAT, PHOSPHOSERINE
1591	1i7w	A	591	812	4.8e-06	0.32	0.96		BETA-CATENIN; CHAIN: A, C; EPTHELIAL-CADHERIN; CHAIN: B, D;	CELL ADHESION E-CADHERIN; E-CADHERIN, CELL ADHESION, BETA-CATENIN, PROTEIN-PROTEIN 2 COMPLEX, EXTENDED INTERFACE, ARMADILLO REPEAT, PHOSPHOSERINE
1591	1i7w	A	676	806	9.6e-05	0.22	0.89		BETA-CATENIN; CHAIN: A, C; EPTHELIAL-CADHERIN; CHAIN: B, D;	CELL ADHESION E-CADHERIN; E-CADHERIN, CELL ADHESION, BETA-CATENIN, PROTEIN-PROTEIN 2 COMPLEX, EXTENDED INTERFACE, ARMADILLO REPEAT, PHOSPHOSERINE
1591	1ial	A	426	806	4.8e-16	0.19	0.43		IMPORTIN ALPHA; CHAIN: A;	NUCLEAR IMPORT RECEPTOR KARYOPHERIN ALPHA; NUCLEAR IMPORT RECEPTOR, NUCLEAR LOCALIZATION SIGNAL, 2 ARMADILLO REPEATS, AUTOINHIBITION, INTRASTERIC REGULATION
1591	1ial	A	733	812	1.2e-06	0.29	0.12		IMPORTIN ALPHA;	NUCLEAR IMPORT RECEPTOR

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A;	KARYOPHERIN ALPHA; NUCLEAR IMPORT RECEPTOR, NUCLEAR LOCALIZATION SIGNAL, 2 ARMADILLO REPEATS, AUTOINHIBITION, INTRASTERIC REGULATION
1591	libr	B	732	806	7.2e-05	-0.29	0.17		RAN; CHAIN: A, C; IMPORTIN BETA SUBUNIT; CHAIN: B, D;	SMALL GTPASE KARYOPHERIN BETA, P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
1591	lihg	A	232	321	9.8e-11	-0.03	0.16		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPIASE IMMUNOPHILIN TETRAPEPTIDE
1591	lihg	A	255	380	8.4e-24	0.00	1.00		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPIASE IMMUNOPHILIN TETRAPEPTIDE
1591	lihg	A	256	353	7.2e-25	0.48	1.00		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPIASE IMMUNOPHILIN TETRAPEPTIDE
1591	3bct		396	801	3.6e-16	0.25	0.83		BETA-CATENIN; CHAIN: NULL;	ARMADILLO REPEAT ARMADILLO REPEAT, BETA-CATENIN, CYTOSKELETON TRANSCRIPTION ZN-FINGER
1592	lfb2	A	1117	1161	0.0048	-0.52	0.21		TRANSCRIPTION FACTOR WSTF; CHAIN: A;	OXIDASE FAD, OXIDASE, D-AMINO ACID
1593	laa8	A	195	518	1.2e-05	0.89	0.06		D-AMINO ACID OXIDASE; CHAIN: A, B;	OXIDOREDUCTASE IMPDH; DEHYDROGENASE, ALPHA-8-BETA-8 BARREL, TIM BARREL, PURINE 2 METABOLISM,
1593	lak5		491	679	0.0024	0.77	-0.13		INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE; CHAIN: NULL;	

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1593	1b3o	A	363	543	0.00048	0.81	-0.05		INOSINE MONOPHOSPHATE DEHYDROGENASE 2; CHAIN: A, B;	OXIDOREDUCTASE, TETRAMER, C4-TETRAMER
1593	1b3o	B	300	545	1.2e-06	0.65	-0.14		INOSINE MONOPHOSPHATE DEHYDROGENASE 2; CHAIN: A, B;	DEHYDROGENASE IMPD, IMPDH; DEHYDROGENASE, IMPD, IMPDH, GUANINE NUCLEOTIDE SYNTHESIS
1593	1b65	A	458	670	3.6e-07	0.39	-0.06		AMINOPEPTIDASE; CHAIN: A, B, C, D, E, F;	DEHYDROGENASE IMPD, IMPDH; HYDROLASE, PEPTIDE DEGRADATION, NTN HYDROLASE
1593	1b8t	A	841	867	1.2e-07	-0.47	0.31		CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
1593	1be3	D	753	852	0.0024	0.40	0.15		CYTOCHROME BC1 COMPLEX; CHAIN: A, B, C, D, E, F, G, H, I, J, K;	ELECTRON TRANSPORT UBIQUINOL CYTOCHROME C OXIDOREDUCTASE, COMPLEX ELECTRON TRANSPORT, CYTOCHROME, MEMBRANE PROTEIN
1593	1cex		141	339	8.4e-15	0.95	0.30		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		174	391	4.8e-14	1.07	-0.15		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		244	455	4.8e-18	0.98	0.45		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		293	526	6e-20	0.80	0.00		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									NULL;	SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		373	565	1.1e-16	0.77	-0.20		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		410	619	2.4e-16	0.48	-0.19		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		419	596	1.1e-16	0.48	-0.19		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		488	668	2.4e-13	0.19	-0.19		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		530	772	3.6e-13	0.24	-0.19		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1cex		578	836	6e-12	0.09	-0.19		CUTINASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1593	1d0s	A	488	818	0.0036	0.57	0.17		NICOTINATE MONONUCLEOTIDE:5-6-CHAIN: A;	TRANSFERASE DINUCLEOTIDE-BINDING MOTIF, PHOSPHORIBOSYL TRANSFERASE
1593	1dbi	A	140	401	6e-12	1.13	-0.15		AK.1 SERINE PROTEASE; CHAIN: A	HYDROLASE HYDROLASE
1593	1dbi	A	216	485	1.1e-10	0.80	-0.12		AK.1 SERINE PROTEASE; CHAIN: A	HYDROLASE HYDROLASE
1593	1dbi	A	347	635	3.6e-12	0.84	-0.17		AK.1 SERINE PROTEASE; CHAIN: A	HYDROLASE HYDROLASE
1593	1dbi	A	411	662	6e-12	0.61	-0.18		AK.1 SERINE PROTEASE; CHAIN: A	HYDROLASE HYDROLASE
1593	1dvj	B	302	463	0.00072	0.88	0.00		OROTIDINE 5'-PHOSPHATE	LYASE OMPDECASE; TTM BARREL, DIMER



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DECARBOXYLASE; CHAIN: A, B, C, D;	
1593	leep	A	330	474	0.00012	0.88	0.59		INOSINE 5-MONOPHOSPHATE DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE ALPHA-BETA BARREL, TIM BARREL, IMPDH, IMP DEHYDROGENASE, 2 LOOP-6, PURINE BIOSYNTHESIS
1593	leua	A	196	393	1.1e-06	0.57	-0.14		KDPG ALDOLASE; CHAIN: A, B, C;	LYASE BETA BARREL, TRIMER, CARBINOLAMINE
1593	leua	A	312	464	2.4e-06	0.77	-0.12		KDPG ALDOLASE; CHAIN: A, B, C;	LYASE BETA BARREL, TRIMER, CARBINOLAMINE
1593	leua	A	321	539	6e-09	1.18	0.34		KDPG ALDOLASE; CHAIN: A, B, C;	LYASE BETA BARREL, TRIMER, CARBINOLAMINE
1593	lfcb	A	375	474	4.8e-05	0.87	-0.08		OXIDOREDUCTASE (CH-OH(D)-CYTOCHROME(A)) FLAVOCYTOCHROME \$B=2= (E.C.1.1.2.3) 1FCB 3	
1593	lfcb	B	303	548	8.4e-18	0.95	0.11		OXIDOREDUCTASE (CH-OH(D)-CYTOCHROME(A)) FLAVOCYTOCHROME \$B=2= (E.C.1.1.2.3) 1FCB 3	
1593	lff9	A	310	515	0.00024	0.60	-0.14		SACCHAROPINE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE LYSINE BIOSYNTHESIS, ALPHA-AMINOADIPATE PATHWAY, 2 SACCHAROPINE REDUCTASE, DEHYDROGENASE
1593	lfch	A	171	798	3.6e-36	0.63	-0.17		PHENOL HYDROXYLASE; CHAIN: A, B, C, D;	FLAVIN FLAVIN, PHENOL HYDROXYLASE, MONOOXYGENASE, OXIDOREDUCTASE
1593	lfch	A	1	619	2.4e-30	0.49	-0.19		PHENOL	FLAVIN FLAVIN, PHENOL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									HYDROXYLASE; CHAIN: A, B, C, D;	HYDROXYLASE, MONOOXYGENASE, OXIDOREDUCTASE
1593	1fch	A	351	824	7.2e-19	0.76	-0.18		PHENOL HYDROXYLASE; CHAIN: A, B, C, D;	FLAVIN FLAVIN, PHENOL HYDROXYLASE, MONOOXYGENASE, OXIDOREDUCTASE
1593	1ga6	A	140	524	7.2e-16	0.48	-0.18		SERINE-CARBOXYL PROTEINASE; CHAIN: A; FRAGMENT OF TYROSTATIN; CHAIN: I;	HYDROLASE PSCP, PSEUDOMONAPEPSIN, PERSTATIN-INSENSITIVE SERINE-CARBOXYL PROTEINASE
1593	1ga6	A	202	549	4.8e-20	0.83	0.40		SERINE-CARBOXYL PROTEINASE; CHAIN: A; FRAGMENT OF TYROSTATIN; CHAIN: I;	HYDROLASE PSCP, PSEUDOMONAPEPSIN, PERSTATIN-INSENSITIVE SERINE-CARBOXYL PROTEINASE
1593	1ga6	A	300	607	4.8e-19	0.75	-0.14		SERINE-CARBOXYL PROTEINASE; CHAIN: A; FRAGMENT OF TYROSTATIN; CHAIN: I;	HYDROLASE PSCP, PSEUDOMONAPEPSIN, PERSTATIN-INSENSITIVE SERINE-CARBOXYL PROTEINASE
1593	1ga6	A	469	819	2.4e-11	0.65	-0.18		SERINE-CARBOXYL PROTEINASE; CHAIN: A; FRAGMENT OF TYROSTATIN; CHAIN: I;	HYDROLASE PSCP, PSEUDOMONAPEPSIN, PERSTATIN-INSENSITIVE SERINE-CARBOXYL PROTEINASE
1593	1geq	A	303	394	0.0024	0.76	-0.09		TRYPTOPHAN SYNTHASE ALPHA-SUBUNIT; CHAIN: A, B;	LYASE TRYPTOPHAN SYNTHASE ALPHA-SUBUNIT, HYPERTHERMOPHILE, 2 PYROCOCOCUS FURIOSUS, X-RAY ANALYSIS, STABILITY, CALORIMETRY
1593	1iml		839	867	8.4e-09	-0.27	0.21		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP, METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
1593	1kap	P	140	481	1.1e-10	0.76	-0.20		ALKALINE PROTEASE; IKAP 4 CHAIN: P; IKAP 5	ZINC METALLOPROTEASE P, ABRUGINOSA ALKALINE

Table 5

SEQ ID NO:	PDB ID	CHAI NID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									TETRAPEPTIDE (GLY SER ASN SER); IKAP 9 CHAIN: 1; IKAP 10	PROTEASE; IKAP 6 CALCIUM BINDING PROTEIN IKAP 19
1593	1kap	P	286	661	2.4e-13	0.73	-0.20		ALKALINE PROTEASE; IKAP 4 CHAIN: P; IKAP 5 TETRAPEPTIDE (GLY SER ASN SER); IKAP 9 CHAIN: 1; IKAP 10	ZINC METALLOPROTEASE P. AERUGINOSA ALKALINE PROTEASE; IKAP 6 CALCIUM BINDING PROTEIN IKAP 19
1593	1lxa		544	644	0.0096	1.66	-0.17		UDP N'-ACETYLGLUCOSAMINE O-ACYLTRANSFERASE; 1LXA 5 CHAIN: NULL; 1LXA 6	ACYLTRANSFERASE 1.PXA; 1LXA 7 TRANSFERASE, ACYLTRANSFERASE, LIPID A BIOSYNTHESIS, 1LXA 17 2 LIPID SYNTHESIS 1LXA 18
1593	1pvd	A	324	515	0.0048	0.79	-0.12		LYASE (CARBON-CARBON) PYRUVATE DECARBOXYLASE (PDC) (E.C.4.1.1.1) 1PVD 3	
1593	1qq4	A	287	432	0.00024	0.68	-0.17		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1593	1qq4	A	338	525	8.4e-09	1.09	-0.05		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1593	1qq4	A	403	616	4.8e-05	0.98	-0.01		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1593	1qq4	A	462	660	1.2e-07	0.82	0.03		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1593	1tal		210	411	1.1e-08	0.59	-0.18		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1593	1tal		299	525	2.4e-10	1.07	-0.20		ALPHA-LYTIC	SERINE PROTEASE SERINE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1593	1tal		416	660	2.4e-08	0.75	-0.19		PROTEASE; CHAIN: NULL;	PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1593	1zfi	A	358	474	0.0024	0.83	0.25		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1593	1zfo		841	866	0.00096	-0.18	0.39		INOSINE MONOPHOSPHATE DEHYDROGENASE; CHAIN: A;	OXIDOREDUCTASE IMPDH, DEHYDROGENASE, CBS DOMAINS
1593	1bqb	A	62	350	0.0081			60.30	LASF-1; CHAIN: NULL;	METAL-BINDING PROTEIN LIM DOMAIN, ZINC-FINGER, METAL-BINDING PROTEIN
1595	1byn	A	18	148	4.2e-49			53.77	AUREOLYSIN; CHAIN: A;	HYDROLASE STAPHYLOCOCCUS AUREUS METALLOPROTEINASE; HYDROLASE, METALLOPROTEINASE
1595	1hs6	A	105	341	2.4e-73	-0.23	0.86		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1595	1hs6	A	15	338	1.4e-62	0.05	0.12		LEUKOTRIENE A-4 HYDROLASE; CHAIN: A;	HYDROLASE PROTEIN-INHIBITOR COMPLEX, ALPHA-BETA PROTEIN
1595	1tsy		13	146	5.6e-49			62.38	LEUKOTRIENE A-4 HYDROLASE; CHAIN: A;	HYDROLASE PROTEIN-INHIBITOR COMPLEX, ALPHA-BETA PROTEIN
1595	1tsy		13	146	5.6e-49			62.38	CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1599	1a25	A	425	558	2.8e-31	0.16	0.82		(CALB) IRSY 3	
									PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALB; CALCIUM++/PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN
1599	1byn	A	425	556	8.4e-38	0.49	1.00		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1599	1c1y	A	440	538	1.1e-14	0.05	-0.03		CYTOSOLIC PHOSPHOLIPASE A2; CHAIN: A, B;	HYDROLASE CPLA2; PHOSPHOLIPASE, LIPID-BINDING, HYDROLASE
1599	1dqv	A	427	621	5.6e-43	0.05	0.70		SYNAPTOTAGMIN III; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS BETA SANDWICH, CALCIUMION, C2 DOMAIN
1599	1dsy	A	425	569	8.4e-32	0.48	0.99		PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	TRANSFERASE CALCIUM++, PHOSPHOLIPID BINDING PROTEIN, CALCIUM-BINDING 2 PROTEIN, PHOSPHATIDYL SERINE, PROTEIN KINASE C
1599	1rsy		421	553	4.2e-38	0.28	1.00		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
1599	3rpb	A	427	561	1.4e-32	0.56	0.74		RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
1600	1ief	I	35	121	1.1e-17	0.08	0.28		CATHERPSIN I, HEAVY	HYDROLASE II FRAGMENT, CD74

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: CHAIN: A, C; CATHEPSIN L: LIGHT CHAIN; CHAIN: B, D; INVARIANT CHAIN; CHAIN: I, J;	FRAGMENT CYSTEINE PROTEINASE, CATHEPSIN, MHC CLASS II, INVARIANT 2 CHAIN, THYROGLOBULIN TYPE-1 DOMAIN
1604	1a6v	L	20	148	7e-30	0.00	-0.17		B1-8; CHAIN: L, H, M, I, N, J;	IMMUNOGLOBULIN IMMUNOGLOBULIN, HAPTEN
1604	1add	L	23	121	8.4e-21	0.06	0.03		IGG4 REA; CHAIN: A; RF- ANIGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIG EN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
1604	1b2w	L	23	70	9.6e-16	0.06	0.80		ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X- RAY STRUCTURE, THREE- DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
1604	1cd0	A	21	143	1.4e-39			51.07	JTO, A VARIABLE DOMAIN FROM LAMBDA-6 TYPE CHAIN: A, B;	IMMUNE SYSTEM IMMUNOGLOBULIN, BENCE- JONES PROTEIN, LAMBDA-6
1604	1cd0	A	23	72	2.4e-25	-0.52	0.41		JTO, A VARIABLE DOMAIN FROM LAMBDA-6 TYPE CHAIN: A, B;	IMMUNE SYSTEM IMMUNOGLOBULIN, BENCE- JONES PROTEIN, LAMBDA-6
1604	1dl7	L	20	148	4.2e-27	0.06	-0.19		ANTIBODY M3C65 (LIGHT CHAIN); CHAIN: L; ANTIBODY M3C65 (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM SINGLE CHAIN FV, REPERTOIRE SHIFT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1604	1dl7	L	20	72	2.4e-29	-0.09	0.01		ANTIBODY M3C65 (LIGHT CHAIN); CHAIN: L; ANTIBODY M3C65 (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM SINGLE CHAIN FV, REPERTOIRE SHIFT
1604	1fvd	A	23	70	1.2e-15	-0.15	0.29		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
1604	1mfa		20	151	7e-30	0.03	-0.19		IMMUNOGLOBULIN FV FRAGMENT (MURINE SE155-4) COMPLEX WITH THE TRISACCHARIDE: 1MFA 3 ALPHA-D-GALACTOSE(1-2)[ALPHA-D-ABEQUOSE(1-3)]ALPHA-1MFA 4 D-MANNOSE (P1-OME) (PART OF THE CELL-SURFACE CARBOHYDRATE 1MFA 5 OF PATHOGENIC SALMONELLA) 1MFA 6	
1604	2cd0	A	23	72	2.4e-25	-0.07	0.28		BENCE-JONES PROTEIN WIL, A VARIABLE DOMAIN FROM CHAIN: A, B;	IMMUNE SYSTEM IMMUNOGLOBULIN, BENCE-JONES PROTEIN, LAMBDA-6
1604	2rhe		15	152	2.8e-45	0.00	-0.08		IMMUNOGLOBULIN BENCE-JONES PROTEIN (LAMBDA, VARIABLE DOMAIN) 2RHE 4	
1604	2rhe		20	161	2.8e-45			67.29	IMMUNOGLOBULIN BENCE-JONES PROTEIN (LAMBDA, VARIABLE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DOMAIN 2RHE 4	
1604	7fab	L	20	170	2.8e-52			57.34	IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
1605	1chc		581	653	0.00096	-0.23	0.07		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
1606	1a4y	A	571	833	7e-13	0.16	0.98		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1606	1a4y	A	623	809	1.1e-15	0.38	0.95		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1606	1a9n	A	624	779	1.2e-05	-0.21	0.34		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1606	1a9n	A	694	810	0.00072	-0.14	0.06		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1606	1brf	A	232	273	0.0084	-0.86	0.21		RUBREDOXIN; CHAIN: A;	ELECTRON TRANSPORT PF RD IRON-SULFUR PROTEIN, HIGH-RESOLUTION STRUCTURE DBREF REMARK
1606	1d0b	A	644	795	1.1e-08	-0.00	0.41		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1606	1ds9	A	701	820	7e-05	-0.12	0.65		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1606	1fo1	A	622	769	0.0029	0.10	0.11		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1606	1fqv	A	556	823	2.8e-36	0.11	0.84		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1606	1fqv	A	764	832	0.00011	0.11	0.12		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1606	1fs1	A	548	597	2.4e-10	-0.46	0.06		CYCLIN A/CDK2-ASSOCIATED P19; CHAIN: A, C; CYCLIN A/CDK2-ASSOCIATED P45; CHAIN: B, D;	LIGASE SKP2 F-BOX, SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1606	1fs1	A	557	597	4.2e-11	-0.32	0.77		CYCLIN A/CDK2-ASSOCIATED P19; CHAIN: A, C; CYCLIN A/CDK2-ASSOCIATED P45; CHAIN: B, D;	LIGASE SKP2 F-BOX; SKP1; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1606	1fs2	A	556	827	2.8e-42	-0.04	0.22		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1606	1fs8	A	622	769	0.0029	-0.04	0.03		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH-REPEAT 2 (LRR) DOMAINS
1606	1fs4		5280	367	2.8e-08	0.04	-0.19		SMOOTH MUSCLE MYOSIN HEAVY CHAIN; CHAIN: S, V; SMOOTH MUSCLE MYOSIN ESSENTIAL LIGHT CHAIN; CHAIN: T, W; SMOOTH MUSCLE MYOSIN REGULATORY LIGHT CHAIN; CHAIN: U, Z;	CONTRACTILE PROTEIN MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSCLE; ELC, MYOSIN ALKALI LIGHT CHAIN; RLQ, MYOSIN REGULATORY LIGHT CHAIN 2, SMOOTH MUSCLE PROTEIN, SMOOTH MUSCLE, MYOSIN SUBFRAGMENT 2, HEAVY 2 MEROMYOSIN, ESSENTIAL LIGHT CHAIN, REGULATORY LIGHT CHAIN, 3 MOTOR PROTEIN, COILED-COIL
1606	1yre	A	616	837	1.4e-06	0.00	0.45		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPL, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMISPHERAL TWINNING, 3 MEROPEDRAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1606	1yr5	A	624	817	2.4e-10	0.15	0.47		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TWINNING, MEROHEDRY TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SP11, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1606	2bnh		571	823	1.3e-16	0.44	1.00		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1606	2bnh		622	797	1.2e-18	0.46	1.00		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1612	1fnf		31	454	1.4e-42			71.47	FIBRONECTIN; 1FNF 6 CHAIN: NULL; 1FNF 7	CELL ADHESION PROTEIN RGD, EXTRACELLULAR MATRIX 1FNF 18
1613	1cvj	A	55	235	1.4e-38			51.11	POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*A P*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM4, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
1613	1ha1		144	238	4.8e-29	0.80	1.00		HNKRN A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNKRN, RBD,

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1613	1ha1		144	279	1.4e-59			110.55	HNRNP A1; CHAIN: NULL;	RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
1613	1ha1		4	91	4.2e-36			72.80	HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
1613	1ha1		53	224	1.4e-59	0.19	0.24		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
1613	2up1	A	143	279	4.2e-62			105.46	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
1613	2up1	A	3	91	4.2e-38			66.13	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1613	2up1	A	52	230	4.2e-62	0.13	-0.01		CHAIN: B; HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
1614	1by2		2	109	2.8e-44			108.34	MAC-2 BINDING PROTEIN; CHAIN: NULL;	EXTRACELLULAR MODULE TUMOR-ASSOCIATED ANTIGEN 90K; EXTRACELLULAR MODULE, SCAVENGER RECEPTOR, TUMOUR-ASSOCIATED 2 ANTIGEN, EXTRACELLULAR MATRIX, GLYCOSYLATED PROTEIN
1614	1by2		38	150	1.4e-44			109.85	MAC-2 BINDING PROTEIN; CHAIN: NULL;	EXTRACELLULAR MODULE TUMOR-ASSOCIATED ANTIGEN 90K; EXTRACELLULAR MODULE, SCAVENGER RECEPTOR, TUMOUR-ASSOCIATED 2 ANTIGEN, EXTRACELLULAR MATRIX, GLYCOSYLATED PROTEIN
1614	1by2		41	146	1.4e-44	1.35	1.00		MAC-2 BINDING PROTEIN; CHAIN: NULL;	EXTRACELLULAR MODULE TUMOR-ASSOCIATED ANTIGEN 90K; EXTRACELLULAR MODULE, SCAVENGER RECEPTOR, TUMOUR-ASSOCIATED 2 ANTIGEN, EXTRACELLULAR MATRIX, GLYCOSYLATED PROTEIN
1614	1f5y	A	2	88	4.2e-23			51.72	LOW-DENSITY LIPOPROTEIN	LIPID BINDING PROTEIN LDL RECEPTOR; BETA HAIRPIN, 3-10

942

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1614	1fyv	A	1	161	5.6e-51			98.94	RECEPTOR; CHAIN: A; TOLL-LIKE RECEPTOR 1; CHAIN: A;	HELIX, CALCIUM BINDING SIGNALING PROTEIN BETA- ALPHA-BETA FOLD PARALLEL BETA SHEET
1614	1fyx	A	5	154	2.8e-44			87.69	TOLL-LIKE RECEPTOR 2; CHAIN: A;	SIGNALING PROTEIN BETA- ALPHA-BETA FOLD
1616	1c1g	A	159	417	7.2e-11	-0.01	0.31		TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
1616	1c1g	A	159	443	7.2e-11			90.49	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
1616	1c1g	A	373	417	0.00056	-0.45	0.12		TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
1616	1dan	L	32	114	5.6e-09	0.01	-0.20		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D- PHE-PHE-ARG- CHLOROMETHYLKETO NE (DFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO- FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
1616	1e7a	A	203	417	1.2e-08	-0.27	0.11		SERUM ALBUMIN; CHAIN: A, B;	CARRIER PROTEIN CARRIER PROTEIN, ALBUMIN, GENERAL ANESTHETIC, PROPOFOL
1616	1f5y	A	30	108	2.8e-12	0.17	-0.14		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A;	LIPID BINDING PROTEIN LDL RECEPTOR; BETA HAIRPIN, 3-10 HELIX, CALCIUM BINDING
1616	1f5y	A	37	107	2.4e-13	0.41	-0.12		LOW-DENSITY LIPOPROTEIN	LIPID BINDING PROTEIN LDL RECEPTOR; BETA HAIRPIN, 3-10

943

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1616	1lhc1	A	159	417	1.2e-10	0.01	0.58		RECEPTOR; CHAIN: A; ALPHA-ACTININ 2; CHAIN: A, B;	HELIX, CALCIUM BINDING TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE- HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN- BINDING PROTEIN
1616	1lhc1	A	321	416	0.00096	0.16	0.18		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE- HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2 Z-LINE, ACTIN- BINDING PROTEIN
1616	1lhy7	A	35	115	7e-09	0.05	-0.20		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL- SURFACE RECEPTOR, CALCIUM- BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
1616	2lrc	P	301	350	0.0011	0.38	-0.07		TRANSDUCIN; CHAIN: B; G; PHOSDUCIN; CHAIN: P;	COMPLEX (TRANSDUCER/TRANSDUCTION) GT BETA-GAMMA; MEKA, PP33; PHOSDUCIN, TRANSDUCIN, BETA-GAMMA, SIGNAL TRANSDUCTION, 2 REGULATON, PHOSPHORYLATION, G PROTEINS, THIOREDOXIN, 3 VISION, MEKA, COMPLEX (TRANSDUCER/TRANSDUCTION)
1618	1stro		352	430	1.4e-22	0.39	1.00		PNPASE; CHAIN: NULL;	S1 RNA-BINDING DOMAIN POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE, S1 RNA-BINDING DOMAIN, POLYNUCLEOTIDE

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1619	1a0j	A	30	226	4.2e-58	-0.08	0.93		TRYPSIN; CHAIN: A, B, C, D;	PHOSPHORYLASE 2 (PNPASE)
1619	1a7s		30	252	2.8e-37			68.47	HEPARIN BINDING PROTEIN; CHAIN: NULL;	SERINE PROTEINASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
1619	1d6w	A	7	122	8.4e-25	-0.20	0.55		THROMBIN; CHAIN: A; DECAPEPTIDE INHIBITOR; CHAIN: I;	SERINE PROTEASE HOMOLOG CAP37, AZUROCIDIN; SERINE PROTEASE HOMOLOG, ENDOTOXIN BINDING, HEPARIN
1619	1ddj	A	12	266	7e-51			59.39	PLASMINOGEN; CHAIN: A, B, C, D;	HYDROLASE/HYDROLASE INHIBITOR HYDROLASE, THROMBIN, THROMBIN INHIBITOR
1619	1dlk	B	30	230	8.4e-53	-0.29	0.53		DELTA-CHYMOTRYPSIN; CHAIN: A, C; DELTA-CHYMOTRYPSIN; CHAIN: B, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC DOMAIN
1619	1elt		30	263	4.2e-49			55.84	ELASTASE; 1ELT 4 CHAIN: NULL; 1ELT 5	HYDROLASE DELTA-CHYMOTRYPSIN, PEPTIDIC INHIBITOR, CHLOROMETHYL KETONE
1619	1elv	A	3	122	2.4e-24	-0.42	0.03		COMPLEMENT C1S COMPONENT; CHAIN: A;	SERINE PROTEINASE
1619	1ept	A	30	71	8.4e-19	-0.51	0.96		HYDROLASE (SERINE PROTEASE) PORCINE E-TRYPSIN (E.C.3.4.21.4) 1EPT 3	HYDROLASE TRYPSIN-LIKE SERIN PROTEASE, CCP (OR SUSHI OR SCR)MODULE
1619	1euf	A	30	261	4.2e-49			59.76	DUODENASE; CHAIN: A;	HYDROLASE BOVINE DUODENASE, SERINE PROTEASE, DUAL SPECIFICITY, 2 CRYSTAL STRUCTURE
1619	1f7z	A	30	226	5.6e-55	-0.05	0.99		TRYPSIN II, ANIONIC;	HYDROLASE/HYDROLASE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1619	1ff8	A	30	221	1.3e-47	-0.15	0.39		CHAIN: A; PANCREATIC TRYPSIN INHIBITOR; CHAIN: I;	INHIBITOR BPTI SERINE PROTEASE, TRYPSIN PRECURSOR
									NATURAL KILLER CELL PROTEASE 1; CHAIN: A, B; ECOTIN; CHAIN: C, E; ECOTIN; CHAIN: D, F;	HYDROLASE/HYDROLASE INHIBITOR GRANZYME B; COMPLEX (SERINE PROTEASE/INHIBITOR), PROTEASE SUBSTRATE 2 INTERACTIONS, BETA STRAND STRUCTURE, CHYMOTRYPSIN FOLD, 3 GRANZYME B, ECOTIN
1619	1fni	A	30	226	1.4e-58	-0.05	0.81		TRYPSIN; CHAIN: A;	HYDROLASE SERINE PROTEASE, HYDROLASE
1619	1fuj	A	30	255	5.6e-46			69.94	PR3; CHAIN: A, B, C, D;	HYDROLASE (SERINE PROTEASE) MYELOBLASTIN; HYDROLASE, SERINE PROTEASE, GLYCOPROTEIN, ZYMOGEN, SIGNAL
1619	1gct	A	30	230	1.1e-52	-0.08	0.21		HYDROLASE (SERINE PROTEINASE) GAMMA-*CHYMOTRYPSIN *A (E.C.3.4.21.1) (\$P*H 7.0) 1GCT 3	
1619	1iau	A	30	266	5.6e-47			57.15	GRANZYME B; CHAIN: A; ACETYL-ISOLEUCYL-GLUTAMYL-PROLYL-ASPARTYL- CHAIN: B;	HYDROLASE ACE-IL-E-GLU-PRO-ASP-CHO; HYDROLASE
1619	1pfx	C	30	241	1.3e-51	0.24	0.55		FACTOR IXA; CHAIN: C, L; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1619	1ppf	E	11	113	1.4e-32			75.05	HYDROLASE(SERINE PROTEINASE) HUMAN LEUKOCYTE ELASTASE (HLE) (NEUTROPHIL ELASTASE (HNE)) IPPF 3 (E.C.3.4.21.37) COMPLEX WITH THE THIRD DOMAIN OF TURKEY IPPF 4 OVOMUCOID INHIBITOR (OMTKY3) IPPF 5	
1619	1ppf	E	30	233	1.3e-45			118.91	HYDROLASE(SERINE PROTEINASE) HUMAN LEUKOCYTE ELASTASE (HLE) (NEUTROPHIL ELASTASE (HNE)) IPPF 3 (E.C.3.4.21.37) COMPLEX WITH THE THIRD DOMAIN OF TURKEY IPPF 4 OVOMUCOID INHIBITOR (OMTKY3) IPPF 5	
1619	1qpf	A	30	266	4.2e-46			55.54	ELASTASE; CHAIN: A;	HYDROLASE (SERINE PROTEASE) PPE; HYDROLASE/SERINE PROTEASE, ATOMIC RESOLUTION
1619	1slw	B	30	226	5.6e-56	0.07	0.71		ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
1619	1tm	A	30	226	4.2e-56	0.10	0.98		HYDROLASE (SERINE	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEINASE) TRYPsin (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR ITRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) ITRN 4 HUMAN TRYPsin, DFP INHIBITED ITRN 6	
1619	2sta	E	30	226	8.4e-60	0.03	0.92		TRYPsin; CHAIN: E; TRYPsin INHIBITOR; CHAIN: I	HYDROLASE/HYDROLASE INHIBITOR SERINE PROTEINASE, TRYPsin INHIBITOR
1619	5ptp		30	227	2.8e-53	-0.00	0.88		BETA TRYPsin; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
1620	1aui	A	1	191	5.6e-86			191.33	SERINE/THREONINE PHOSPHATASE 2B; CHAIN: A, B;	HYDROLASE CALCINEURIN; HYDROLASE, PHOSPHATASE, IMMUNOSUPPRESSION
1620	1aui	A	1	291	2.8e-99			289.74	SERINE/THREONINE PHOSPHATASE 2B; CHAIN: A, B;	HYDROLASE CALCINEURIN; HYDROLASE, PHOSPHATASE, IMMUNOSUPPRESSION
1620	1aui	A	41	272	2.8e-99	0.41	1.00		SERINE/THREONINE PHOSPHATASE 2B; CHAIN: A, B;	HYDROLASE CALCINEURIN; HYDROLASE, PHOSPHATASE, IMMUNOSUPPRESSION
1620	1tco	A	1	191	5.6e-86			204.13	SERINE/THREONINE PHOSPHATASE B2; CHAIN: A, B; FK506-BINDING PROTEIN; CHAIN: C;	COMPLEX (HYDROLASE/SOMERASE) CALCINEURIN A, CAM-PRP CATALYTIC SUBUNIT, FKBP-12, CIS-TRANS ISOMERASE; COMPLEX (HYDROLASE/SOMERASE), IMMUNOSUPPRESSANT
1620	1tco	A	1	267	2.8e-98			290.51	SERINE/THREONINE PHOSPHATASE B2;	COMPLEX (HYDROLASE/SOMERASE)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A, B; FK506-BINDING PROTEIN; CHAIN: C;	CALCINEURIN A, CAM-PRP CATALYTIC SUBUNIT, FKBP-12, CIS-TRANS ISOMERASE; COMPLEX (HYDROLASE/ISOMERASE), IMMUNOSUPPRESSANT
1620	ltco	A	41	267	2.8e-98	0.31	1.00		SERINE/THREONINE PHOSPHATASE B2; CHAIN: A, B; FK506-BINDING PROTEIN; CHAIN: C;	COMPLEX (HYDROLASE/ISOMERASE) CALCINEURIN A, CAM-PRP CATALYTIC SUBUNIT, FKBP-12, CIS-TRANS ISOMERASE; COMPLEX (HYDROLASE/ISOMERASE), IMMUNOSUPPRESSANT
									GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE GAPDH; GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE, GLYCOSOME, 2 TRYPTANOSOME, OXIDOREDUCTASE
1625	la7k	A	62	338	4.2e-25	-0.17	0.01		GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE GAPDH; GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE, GLYCOSOME, 2 TRYPTANOSOME, OXIDOREDUCTASE
1625	lbrm	A	1	192	7e-63			167.88	ASPARTATE-SEMIALDEHYDE DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE ASADH, ASDH; CRYSTAL STRUCTURE, DEHYDROGENASE, ESCHERICHIA COLI, ENZYME, 2 NADP
1625	lbrm	A	2	308	0	0.50	1.00		ASPARTATE-SEMIALDEHYDE DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE ASADH, ASDH; CRYSTAL STRUCTURE, DEHYDROGENASE, ESCHERICHIA COLI, ENZYME, 2 NADP
									ASPARTATE-SEMIALDEHYDE DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE ASADH, ASDH; CRYSTAL STRUCTURE, DEHYDROGENASE, ESCHERICHIA COLI, ENZYME, 2 NADP
1625	lbrm	A	2	372	0			449.11	ASPARTATE-SEMIALDEHYDE DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE ASADH, ASDH; CRYSTAL STRUCTURE, DEHYDROGENASE, ESCHERICHIA COLI, ENZYME, 2 NADP
									ASPARTATE-SEMIALDEHYDE DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE ASADH, ASDH; CRYSTAL STRUCTURE, DEHYDROGENASE, ESCHERICHIA COLI, ENZYME, 2 NADP
1625	lbrm	A	3	129	1.4e-57			58.34	ASPARTATE-SEMIALDEHYDE DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE ASADH, ASDH; CRYSTAL STRUCTURE, DEHYDROGENASE, ESCHERICHIA COLI, ENZYME, 2 NADP

949

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1626	1b7y	B	7	170	2.8e-43	0.10	1.00		PHENYLALANYL-TRNA SYNTHETASE; CHAIN: A; PHENYLALANYL-TRNA SYNTHETASE; CHAIN: B;	LIGASE PHERS; PHERS; ENZYME, TRNA SYNTHETASE, ALPHA/BETA HOMODIMER
1628	1b9m	A	28	97	0.00011	-0.87	0.19		MODE; CHAIN: A, B;	TRANSCRIPTION DNA-BINDING, GENE REGULATION, WINGED HELIX TURN HELIX, 2 MOL YBDATE, OB FOLD
1628	1b9m	A	8	97	0.00056	0.13	0.36		MODE; CHAIN: A, B;	TRANSCRIPTION DNA-BINDING, GENE REGULATION, WINGED HELIX TURN HELIX, 2 MOL YBDATE, OB FOLD
1628	1b12	B	10	86	0.00014	-0.09	0.09		DIPHThERIA TOXIN REPRESSOR; CHAIN: A, B;	REPRESSOR DTXR; REPRESSOR, TRANSCRIPTION REGULATION, DNA-BINDING, IRON
1628	1c0w	A	10	86	0.00014	-0.21	0.07		DIPHThERIA TOXIN REPRESSOR; CHAIN: A, B, C, D; DNA (5'-CHAIN: E; DNA (3'-CHAIN: F;	GENE REGULATION/DNA IRON-DEPENDENT DIPHThERIA TOX REGULATORY ELEMENT, TOXIN REPRESSOR-DNA COMPLEX, METAL BINDING SH-LIKE DOMAIN
1628	1ddn	A	10	86	0.00014	0.12	0.06		DIPHThERIA TOX REPRESSOR; CHAIN: A, B, C, D; DIPHThERIA TOX DNA OPERATOR; CHAIN: E, F;	COMPLEX (REGULATORY PROTEIN/DNA) DTXR; TOXPO; COMPLEX (REGULATORY PROTEIN/DNA), DIPHThERIA TOX REPRESSOR, 2 TRANSCRIPTION REGULATION, DNA-BINDING REGULATORY PROTEIN, 3 IRON-REGULATED REPRESSOR, DNA-PROTEIN, DNA-REPRESSOR 4 COMPLEX
1628	2dtr		10	86	0.00014	-0.05	0.22		DIPHThERIA TOXIN REPRESSOR; CHAIN: NULL;	REPRESSOR DTXR; TRANSCRIPTION REGULATION, REPRESSOR, DNA-BINDING, IRON

950

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1631	Inal	1	21	139	0.0048	-0.03	0.03		N-ACETYLNEURAMINATE LYASE; INAL 4 CHAIN: 1, 2, 3, 4; INAL 5	LYASE
1631	lzym	A	1	57	1.4e-12	-0.79	1.00		ENZYME I; CHAIN: A, B;	PHOSPHOTRANSFERASE PHOSPHOTRANSFERASE
1632	la1s		263	599	0	0.20	0.86		ORNITHINE CARBAMOYLTRANSFERASE; CHAIN: NULL;	TRANSCARBAMYLASE ORNITHINE TRANSCARBAMYLASE; TRANSCARBAMYLASE
1632	la1s		289	601	0			78.86	ORNITHINE CARBAMOYLTRANSFERASE; CHAIN: NULL;	TRANSCARBAMYLASE ORNITHINE TRANSCARBAMYLASE; TRANSCARBAMYLASE
1632	lduv	G	262	597	9.8e-85	0.08	0.31		ORNITHINE TRANSCARBAMOYLASE; CHAIN: G, H, I;	TRANSFERASE OCTASE-1; ENZYME-INHIBITOR COMPLEX
1632	ldxh	A	263	599	1.1e-85	-0.05	0.21		ORNITHINE CARBAMOYLTRANSFERASE; CHAIN: A;	TRANSFERASE ORNITHINE TRANSCARBAMYLASE; TRANSFERASE, TRANSCARBAMYLASE
1632	ldxh	A	402	585	9.6e-24	0.50	1.00		ORNITHINE CARBAMOYLTRANSFERASE; CHAIN: A;	TRANSFERASE ORNITHINE TRANSCARBAMYLASE; TRANSFERASE, TRANSCARBAMYLASE
1632	loth	A	263	609	2.8e-93	-0.00	0.70		ORNITHINE TRANSCARBAMOYLASE; CHAIN: A;	TRANSFERASE TRANSCARBAMOYLASE, TRANSFERASE
1632	loth	A	270	608	2.8e-93			75.53	ORNITHINE TRANSCARBAMOYLASE; CHAIN: A;	TRANSFERASE TRANSCARBAMOYLASE, TRANSFERASE
1632	zalc	A	262	598	4.2e-57	0.04	-0.09		TRANSFERASE (CARBAMOYL-P-ASPARTATE)	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ASPARTATE CARBAMOYLTRANSFERASE (ASPARTATE TRANSFERASE) 2ATC 4 (E.C.2.1.3.2) 2ATC 5	
1632	3csu	A	262	598	8.4e-61	0.21	0.11		ASPARTATE CARBAMOYLTRANSFERASE; CHAIN: A, B, C;	TRANSFERASE (CARBAMOYL-P, ASPARTATE) TRANSFERASE (CARBAMOYL-P, ASPARTATE)
1632	3csu	C	262	598	2.8e-67	-0.05	0.04		ASPARTATE CARBAMOYLTRANSFERASE; CHAIN: A, B, C;	TRANSFERASE (CARBAMOYL-P, ASPARTATE) TRANSFERASE (CARBAMOYL-P, ASPARTATE)
1634	1bk7	A	24	202	2.4e-16			54.70	RIBONUCLEASE MCI; CHAIN: A;	HYDROLASE RNASE MCI; HYDROLASE (NUCLEIC ACID, RNA)
1634	1bk7	A	78	138	2.4e-16	-0.11	0.06		RIBONUCLEASE MCI; CHAIN: A;	HYDROLASE RNASE MCI; HYDROLASE (NUCLEIC ACID, RNA)
1634	1dix	A	10	206	3.6e-17			59.98	EXTRACELLULAR RIBONUCLEASE LE; CHAIN: A;	HYDROLASE ALPHA PLUS BETA
1634	1dix	A	78	138	3.6e-17	0.13	0.55		EXTRACELLULAR RIBONUCLEASE LE; CHAIN: A;	HYDROLASE ALPHA PLUS BETA
1635	1c1g	A	70	346	7.2e-09			64.80	TROPOMYOSIN; CHAIN: A, B, C, D	CONTRACTILE PROTEIN TROPOMYOSIN COILED-COIL ALPHA-HELICAL, CONTRACTILE PROTEIN
1635	1hcl	A	126	274	4.8e-09	-0.27	0.10		ALPHA-ACTININ 2; CHAIN: A, B;	TRIPLE-HELIX COILED COIL ALPHA ACTININ SKELETAL MUSCLE ISOFORM 2, TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN, MUSCLE, 2-Z-LINE, ACTIN-

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1638	1br1	B	32	176	1.2e-29			73.41	MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	BINDING PROTEIN MUSCLE PROTEIN MADE; MUSCLE PROTEIN
1638	1br1	B	38	173	1.2e-29	0.16	1.00		MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MADE; MUSCLE PROTEIN
1638	1cll		30	174	6e-35			111.80	CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
1638	1cll		38	173	6e-35	0.06	1.00		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
1638	1cmf		1	66	7e-34			51.56	CALMODULIN (VERTEBRATE); 1CMF 6 CHAIN: NULL; 1CMF 7	CALCIUM-BINDING PROTEIN CALMODULIN APO TR2C-DOMAIN; 1CMF 9
1638	1ext	A	28	173	1.2e-34			107.65	CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
1638	1ext	A	38	173	1.2e-34	0.30	1.00		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
1638	1fw4	A	1	63	1.4e-32			53.23	CALMODULIN; CHAIN: A;	METAL BINDING PROTEIN EF-HAND, HELIX-LOOP-HELIX, FRAGMENT, CALCIUM, TR2C, C-2 TERMINAL DOMAIN, CALMODULIN
1638	1ggw	A	33	175	2.4e-08			73.80	CDC4P; CHAIN: A;	CYTOKINE EF-HAND PROTEIN, MYOSIN LIGHT CHAIN, LIGHT CHAIN, CYTOKINESIS, CELL CYCLE, EF-HAND
1638	1jfo	A	31	111	2.4e-11	0.61	0.95		OBELIN; CHAIN: A;	LUMINESCENT PROTEIN BIOLUMINESCENCE, CALCIUM-REGULATED PHOTOPROTEIN, OBELIN, 2 OBELLA, HYDROID
1638	1tnc	A	1	65	2.8e-33			52.26	CALCIUM BINDING	



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEIN CALMODULIN (TR=2=C\$ FRAGMENT COMPRISING RESIDUES 78 - 148 ITRC 3 OF THE INTACT MOLECULE) ITRC 4	
1638	lvtk	A	27	175	1.2e-34			110.08	CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
1638	lvtk	A	38	173	1.2e-34	0.18	1.00		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
1643	ljf2	R	197	276	1.1e-32	0.53	0.99		23S RRNA; CHAIN: G; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L6; CHAIN: E; RIBOSOMAL PROTEIN L7AE; CHAIN: F; RIBOSOMAL PROTEIN L10; CHAIN: G; RIBOSOMAL PROTEIN L10E; CHAIN: H; RIBOSOMAL PROTEIN	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10; 50S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN P0, HMAL10, L10E; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									L13; CHAIN: I; RIBOSOMAL PROTEIN L14; CHAIN: J; RIBOSOMAL PROTEIN L15; CHAIN: K; RIBOSOMAL PROTEIN L15E; CHAIN: L; RIBOSOMAL PROTEIN L18; CHAIN: M; RIBOSOMAL PROTEIN L18E; CHAIN: N; RIBOSOMAL PROTEIN L19E; CHAIN: O; RIBOSOMAL PROTEIN L21E; CHAIN: P; RIBOSOMAL PROTEIN L22; CHAIN: Q; RIBOSOMAL PROTEIN L23; CHAIN: R; RIBOSOMAL PROTEIN L24; CHAIN: S; RIBOSOMAL PROTEIN L24E; CHAIN: T; RIBOSOMAL PROTEIN L29; CHAIN: U; RIBOSOMAL PROTEIN L30; CHAIN: V; RIBOSOMAL PROTEIN L31E; CHAIN: W; RIBOSOMAL PROTEIN L32E; CHAIN: X; RIBOSOMAL PROTEIN L37AE; CHAIN: Y; RIBOSOMAL PROTEIN L37E; CHAIN: Z;	HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN-PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1643	1jj2	R	197	276	5.6e-27	0.53	0.99		RIBOSOMAL PROTEIN L39E; CHAIN: I; RIBOSOMAL PROTEIN L44E; CHAIN: 2; 23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L6; CHAIN: E; RIBOSOMAL PROTEIN L7A; CHAIN: F; RIBOSOMAL PROTEIN L10; CHAIN: G; RIBOSOMAL PROTEIN L10E; CHAIN: H; RIBOSOMAL PROTEIN L13; CHAIN: I; RIBOSOMAL PROTEIN L14; CHAIN: J; RIBOSOMAL PROTEIN L15; CHAIN: K; RIBOSOMAL PROTEIN L15E; CHAIN: L; RIBOSOMAL PROTEIN L18; CHAIN: M; RIBOSOMAL PROTEIN L18E; CHAIN: N; RIBOSOMAL PROTEIN	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4, 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10; 50S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN P0, HMAL10, L10E; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1644	1q44	A	141	313	8.4e-08	0.67	-0.11		L19E: CHAIN: O; RIBOSOMAL PROTEIN L21E; CHAIN: P; RIBOSOMAL PROTEIN L22; CHAIN: Q; RIBOSOMAL PROTEIN L23; CHAIN: R; RIBOSOMAL PROTEIN L24; CHAIN: S; RIBOSOMAL PROTEIN L24E; CHAIN: T; RIBOSOMAL PROTEIN L29; CHAIN: U; RIBOSOMAL PROTEIN L30; CHAIN: V; RIBOSOMAL PROTEIN L31E; CHAIN: W; RIBOSOMAL PROTEIN L32E; CHAIN: X; RIBOSOMAL PROTEIN L37AE; CHAIN: Y; RIBOSOMAL PROTEIN L37E; CHAIN: Z; RIBOSOMAL PROTEIN L39E; CHAIN: 1; RIBOSOMAL PROTEIN L44E; CHAIN: 2;	RIBOSOMAL PROTEIN L31E, L34, HI.30; 50S RIBOSOMAL PROTEIN L32E, HI.5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HI.39E, HI.46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN-PROTEIN
1644	1q44	A	185	330	1.2e-06	0.46	-0.09		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1645	1ehd	A	680	767	2.8e-05	-0.01	0.06		AGGLUTININ ISOLLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1645	1emn		490	555	2.8e-09	0.03	-0.20		FIBRILLIN; CHAIN: NULL;	DOMAINS MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1645	1emn		525	631	1.4e-07	0.62	-0.07		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1645	1emn		677	747	7e-09	0.19	-0.18		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1645	1fak	L	360	443	2.8e-09	0.02	-0.20		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; SLI5; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1645	1g44	A	570	755	2.4e-08	0.16	-0.19		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1645	1g44	C	404	630	4.8e-08	0.07	-0.17		COMPLEMENT CONTROL PROTEIN; CHAIN: A, B, C;	IMMUNE SYSTEM BETA, MODULE
1645	1igr	A	408	679	6e-09	0.33	-0.19		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
1645	1k1o		363	523	2.8e-11	0.11	-0.19		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1645	1k1o		403	555	5.6e-09	0.07	-0.20		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1645	1k1o		469	624	1.4e-09	0.19	-0.18		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1645	1k1o		532	676	4.2e-10	0.30	-0.17		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
1645	1slz		352	456	1.4e-08	0.08	-0.20		ANTISTASIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR FACTOR XA INHIBITOR; ANTISTASIN, CRYSTAL STRUCTURE, FACTOR XA INHIBITOR, 2 SERINE PROTEASE INHIBITOR, THROMBOSIS
1645	4mt2		434	488	5.6e-08	-0.00	-0.19		METALLOTHIONEIN METALLOTHIONEIN ISOFORM II 4MT2.3	
1645	9wga	A	337	497	1.4e-18	0.36	-0.14		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	
1645	9wga	A	400	560	5.6e-18	0.24	-0.18		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	
1645	9wga	A	445	612	1.4e-17	0.28	-0.17		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA.3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1645	9wga	A	460	673	1.3e-16	0.30	-0.09		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1645	9wga	A	519	706	2.8e-15	0.16	-0.14		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1645	9wga	A	584	742	2.8e-17	0.47	-0.13		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1645	9wga	A	622	762	4.2e-15	0.26	-0.12		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1650	1ck4	A	8	253	5.6e-49			170.57	INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I- DOMAIN, METAL BINDING, COLLAGEN, ADHESION
1651	1aox	A	5	258	1.4e-44			145.63	INTEGRIN ALPHA 2 BETA; CHAIN: A, B;	INTEGRIN INTEGRIN, CELL ADHESION, GLYCOPROTEIN
1651	1aox	A	83	337	8.4e-44			146.30	INTEGRIN ALPHA 2 BETA; CHAIN: A, B;	INTEGRIN INTEGRIN, CELL ADHESION, GLYCOPROTEIN
1651	1aox	A	84	335	8.4e-44	0.53	1.00		INTEGRIN ALPHA 2 BETA; CHAIN: A, B;	INTEGRIN INTEGRIN, CELL ADHESION, GLYCOPROTEIN
1651	1atz	A	272	331	2.4e-09	-0.39	0.07		VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN- BINDING, HEMOSTASIS, DINUCLEOTIDE BINDING FOLD
1651	1atz	A	86	268	3.6e-23			57.40	VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN- BINDING, HEMOSTASIS, DINUCLEOTIDE BINDING FOLD
1651	1atz	A	88	175	3.6e-23	0.35	0.93		VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN- BINDING, HEMOSTASIS, DINUCLEOTIDE BINDING FOLD

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1651	1atz	A	88	306	1.3e-16	0.07	0.63		VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN-BINDING, HEMOSTASIS, DINUCLEOTIDE BINDING FOLD
1651	1atz	A	8	191	2.8e-15			55.88	VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN-BINDING, HEMOSTASIS, DINUCLEOTIDE BINDING FOLD
1651	1auq		1	228	4.2e-22			51.56	A1 DOMAIN OF VON WILLEBRAND FACTOR; CHAIN: NULL;	WILLEBRAND WILLEBRAND, BLOOD COAGULATION, PLATELET, GLYCOPROTEIN
1651	1auq		69	306	7e-22			56.38	A1 DOMAIN OF VON WILLEBRAND FACTOR; CHAIN: NULL;	WILLEBRAND WILLEBRAND, BLOOD COAGULATION, PLATELET, GLYCOPROTEIN
1651	1auq		84	342	7e-22	0.06	-0.02		A1 DOMAIN OF VON WILLEBRAND FACTOR; CHAIN: NULL;	WILLEBRAND WILLEBRAND, BLOOD COAGULATION, PLATELET, GLYCOPROTEIN
1651	1ck4	A	86	331	2.8e-48			170.36	INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
1651	1ck4	A	87	332	2.8e-48	0.71	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
1651	1dzi	A	244	323	1.2e-16	0.15	1.00		INTEGRIN; CHAIN: A; COLLAGEN; CHAIN: B, C, D;	INTEGRIN INTEGRIN, COLLAGEN
1651	1dzi	A	86	324	1.1e-38			120.10	INTEGRIN; CHAIN: A; COLLAGEN; CHAIN: B, C, D;	INTEGRIN INTEGRIN, COLLAGEN
1651	1dzi	A	88	174	1.2e-19	0.35	1.00		INTEGRIN; CHAIN: A; COLLAGEN; CHAIN: B, C, D;	INTEGRIN INTEGRIN, COLLAGEN
1651	1dzi	A	88	319	1.1e-38	0.40	1.00		INTEGRIN; CHAIN: A; COLLAGEN; CHAIN: B, C, D;	INTEGRIN INTEGRIN, COLLAGEN
1651	1dzi	A	8	246	2.8e-38			118.42	INTEGRIN; CHAIN: A;	INTEGRIN INTEGRIN, COLLAGEN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									COLLAGEN; CHAIN: B, C, D;	
1651	1fts	A	233	337	7.2e-13	0.06	0.10		IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
1651	1fts	A	6	197	2.8e-22			50.42	IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
1651	1fts	A	84	205	8.4e-22	0.42	0.46		IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
1651	1fts	A	84	275	8.4e-22			52.21	IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
1651	1ido		10	251	2.8e-38			55.95	INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1651	1ido		88	329	5.6e-34			55.68	INTEGRIN; CHAIN: NULL;	CYTOSKELETON
1651	1ido		89	329	5.6e-34	0.45	1.00		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
1651	1ido		89	329	5.6e-34	0.45	1.00		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
1651	1lfa	A	88	334	7e-30	0.35	0.69		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-1, BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
1651	1qc5	A	7	252	4.2e-48			163.00	ALPHA1 BETA1 INTEGRIN; CHAIN: A; ALPHA1 BETA1 INTEGRIN; CHAIN: B;	CELL ADHESION INTEGRIN, CELL ADHESION
1651	1qc5	A	85	330	9.8e-47	0.47	1.00		ALPHA1 BETA1 INTEGRIN; CHAIN: A; ALPHA1 BETA1 INTEGRIN; CHAIN: B;	CELL ADHESION INTEGRIN, CELL ADHESION
1651	1qc5	A	85	330	9.8e-47			163.53	ALPHA1 BETA1 INTEGRIN; CHAIN: A; ALPHA1 BETA1 INTEGRIN; CHAIN: B;	CELL ADHESION INTEGRIN, CELL ADHESION
1652	1a14	H	1	93	1.4e-46			62.07	NEURAMINIDASE; CHAIN: N; SINGLE CHAIN ANTIBODY; CHAIN: H, L;	COMPLEX (ANTIBODY/ANTIGEN) COMPLEX (ANTIBODY/ANTIGEN), SINGLE-CHAIN ANTIBODY, 2 GLYCOSYLATED PROTEIN
1652	1a6v	H	1	94	1.4e-46			68.99	B1-8; CHAIN: L, H, M, I, N, J;	IMMUNOGLOBULIN IMMUNOGLOBULIN, HAPTEN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1652	1ayl	H	206	315	1.4e-55	0.36	0.84		TP7 FAB; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, ANTIBODY, FAB, ENZYME INHIBITOR, PCR, 2 HOT START
1652	1ayl	H	20	232	5.6e-55			150.59	TP7 FAB; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, ANTIBODY, FAB, ENZYME INHIBITOR, PCR, 2 HOT START
1652	1ayl	H	21	141	5.6e-55	0.70	1.00		TP7 FAB; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, ANTIBODY, FAB, ENZYME INHIBITOR, PCR, 2 HOT START
1652	1bfo	B	21	172	8.4e-35	0.44	0.69		CAMPATH-1G ANTIBODY; CHAIN: A, B, C, D, E, F, G, H;	ANTIBODY ANTIBODY, FAB, CAMPATH-1G, CD52
1652	1bvk	B	1	103	5.6e-46			113.70	HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-L-YSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
1652	1bvk	B	206	316	1.4e-56	0.56	0.94		HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-L-YSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
1652	1bvk	B	20	135	1.3e-54			125.75	HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-L-YSOZYME, 2 COMPLEX (HUMANIZED

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1652	1bvk	B	20	141	1.3e-54	0.59	1.00		HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	ANTIBODY/HYDROLASE) COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
1652	1c12	B	20	231	1.4e-51			139.41	ANTIBODY FRAGMENT FAB; CHAIN: A; ANTIBODY FRAGMENT FAB; CHAIN: B;	IMMUNE SYSTEM ANTIBODY- ANTIGEN COMPLEX, SCFV FRAGMENT, CDRH3, MUSK 2 ODORANT, ODORANT SPECIFICITY, IMMUNE SYSTEM
1652	1c5d	B	20	257	7e-47			127.26	MONOCLONAL ANTIBODY AGAINST THE MAIN IMMUNOGENIC CHAIN: L, A; MONOCLONAL ANTIBODY AGAINST THE MAIN IMMUNOGENIC CHAIN: H, B;	IMMUNE SYSTEM IMMUNOGLOBULIN
1652	1c5d	B	21	172	7e-47	0.44	1.00		MONOCLONAL ANTIBODY AGAINST THE MAIN IMMUNOGENIC CHAIN: L, A; MONOCLONAL ANTIBODY AGAINST THE MAIN IMMUNOGENIC CHAIN: H, B;	IMMUNE SYSTEM IMMUNOGLOBULIN
1652	1ce1	H	206	316	4.2e-55	0.31	0.81		CAMPATH-1H; LIGHT CHAIN; CHAIN: L; CAMPATH-1H; HEAVY CHAIN; CHAIN: H;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1652	1ce1	H	20	144	2.8e-54	0.81	1.00		PEPTIDE ANTIGEN; CHAIN: P; CAMPATH-1H:LIGHT CHAIN; CHAIN: L; CAMPATH-1H:HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
1652	1ce1	H	20	224	2.8e-54			137.30	CAMPATH-1H:LIGHT CHAIN; CHAIN: L; CAMPATH-1H:HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
1652	1cf8	H	207	316	5.6e-55	0.59	0.94		CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;	CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;
1652	1cf8	H	21	147	2.8e-55	0.49	1.00		CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;	CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;
1652	1cf7	H	1	93	2.8e-37			71.49	IGG1 ANTIBODY 1696 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 1696 (VARIABLE HEAVY CHAIN); CHAIN: H; IGG1 ANTIBODY 1696 (CONSTANT HEAVY CHAIN); CHAIN: L;	IMMUNE SYSTEM IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1; IMMUNOGLOBULIN, IGG1 FAB FRAGMENT, CROSS-REACTIVITY, HIV1 PROTEASE, ENZYME 2 INHIBITION, IMMUNOGLOBULIN
1652	1clz	H	21	172	1.3e-33	0.38	0.34		IGG FAB (IGG3, KAPPA);	IMMUNOGLOBULIN MBR96 FAB

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: L, H;	(IMMUNOGLOBULIN); IMMUNOGLOBULIN C REGION, GLYCOPROTEIN, TRANSMEMBRANE
1652	1dby	A	12	122	1.4e-32			69.28	CHLOROPLAST THIOREDOXIN M.CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
1652	1dm0	B	206	316	1.4e-64	0.48	0.98		IGM-KAPPA COLD AGGLUTININ (LIGHT CHAIN); CHAIN: A, C; IGM-KAPPA COLD AGGLUTININ (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM FAB, IGM, ANTIBODY, COLD AGGLUTININ, HUMAN
1652	1dm0	B	20	240	5.6e-62			141.50	IGM-KAPPA COLD AGGLUTININ (LIGHT CHAIN); CHAIN: A, C; IGM-KAPPA COLD AGGLUTININ (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM FAB, IGM, ANTIBODY, COLD AGGLUTININ, HUMAN
1652	1dm0	B	21	144	5.6e-62	0.75	1.00		IGM-KAPPA COLD AGGLUTININ (LIGHT CHAIN); CHAIN: A, C; IGM-KAPPA COLD AGGLUTININ (HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM FAB, IGM, ANTIBODY, COLD AGGLUTININ, HUMAN
1652	1dqd	H	20	246	1.4e-51			131.91	FAB HGR-2 F6; CHAIN: L; FAB HGR-2 F6; CHAIN: H;	IMMUNE SYSTEM GLUCAGON RECEPTOR, MONOCLONAL ANTIBODY, FAB, RECEPTOR 2 ANTAGONIST, TYPICAL IMMUNOGLOBULIN FOLD, LIGHT CHAIN, HEAVY 3 CHAIN, ANTIGEN BINDING SITE, COMPLEMENTARITY- DETERMINING 4 REGIONS

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1652	1e08	H	20	258	1.2e-49	-0.09	0.21		HEMAGGLUTININ (HA1 CHAIN); CHAIN: A; HEMAGGLUTININ (HA2 CHAIN); CHAIN: B; ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	VIRUS/VIRAL PROTEIN COMPLEX (HEMAGGLUTININ/IMMUNOGLOBULIN), HEMAGGLUTININ, 2 IMMUNOGLOBULIN, VIRAL PROTEIN, IMMUNE SYSTEM COMPLEX
1652	1ezv	X	20	124	2.4e-64	0.69	1.00		UBIQUINOL- CYTOCHROME C REDUCTASE COMPLEX CORE CHAIN: A; UBIQUINOL- CYTOCHROME C REDUCTASE COMPLEX CORE CHAIN: B; CYTOCHROME B; CHAIN: C; CYTOCHROME C1; CHAIN: D; UBIQUINOL- CYTOCHROME C REDUCTASE IRON-SULFUR CHAIN: E; UBIQUINOL- CYTOCHROME C REDUCTASE COMPLEX 17 KD CHAIN: H; UBIQUINOL- CYTOCHROME C REDUCTASE COMPLEX 14 KD CHAIN: F; UBIQUINOL- CYTOCHROME C REDUCTASE COMPLEX CHAIN: G; UBIQUINOL-	OXIDOREDUCTASE/ELECTRON TRANSPORT CYTOCHROME BC1 COMPLEX, COMPLEX III, QCR, MITOCHONDRIA, 2 YEAST, ANTIBODY FV-FRAGMENT, STIGMATELLIN, COENZYME Q6, 3 MATRIX PROCESSING PEPTIDASES, UBIQUINONE, ELECTRON TRANSFER, 4 PROTON TRANSFER, Q-CYCLE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CYTOCHROME C REDUCTASE COMPLEX 7.3 KD CHAIN: I; HEAVY CHAIN (VH) OF FV- FRAGMENT; CHAIN: X; LIGHT CHAIN (VL) OF FV-FRAGMENT; CHAIN: Y;	
1652	1f11	B	20	258	3.6e-50	0.08	0.18		F124 IMMUNOGLOBULIN (KAPPA LIGHT CHAIN); CHAIN: A, C; F124 IMMUNOGLOBULIN (GG1 HEAVY CHAIN); CHAIN: B, D;	IMMUNE SYSTEM IMMUNOGLOBULIN, ANTIBODY, FAB, HEPATITIS B, PRES2
1652	1f58	H	207	316	1.4e-54	0.43	0.98		IGG1 ANTIBODY 58.2 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 58.2 (HEAVY CHAIN); CHAIN: H; EXTERIOR MEMBRANE GLYCOPROTEIN(GP120); CHAIN: P;	IMMUNE SYSTEM FAB 58.2; FAB 58.2; V3 LOOP; IMMUNOGLOBULIN, FAB, HIV-1, GP120, V3, IMMUNE SYSTEM
1652	1f58	H	21	147	7e-55	0.58	1.00		IGG1 ANTIBODY 58.2 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 58.2 (HEAVY CHAIN); CHAIN: H; EXTERIOR MEMBRANE GLYCOPROTEIN(GP120); CHAIN: P;	IMMUNE SYSTEM FAB 58.2; FAB 58.2; V3 LOOP; IMMUNOGLOBULIN, FAB, HIV-1, GP120, V3, IMMUNE SYSTEM
1652	1f8f	H	207	316	4.2e-57	0.33	1.00		ANTIBODY FAB FRAGMENT (LIGHT CHAIN); CHAIN: L; ANTIBODY FAB	IMMUNE SYSTEM MONOCLONAL ANTIBODY, ANTIGEN-BINDING FRAGMENT, INTERLEUKIN- 2 2, X-RAY ANALYSIS, CRYSTAL



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FRAGMENT (HEAVY CHAIN); CHAIN: H	
1652	1f8t	H	21	147	4.2e-56	0.68	1.00		ANTIBODY FAB FRAGMENT (LIGHT CHAIN); CHAIN: L; ANTIBODY FAB FRAGMENT (HEAVY CHAIN); CHAIN: H	IMMUNE SYSTEM MONOCLONAL ANTIBODY, ANTIGEN-BINDING FRAGMENT, INTERLEUKIN-2 2, X-RAY ANALYSIS, CRYSTAL
1652	1fb6	A	14	121	7e-33			68.84	THIOREDOXIN M <sub>2</sub> CHAIN: A, B; MUTANT AL2 6E7S9G; CHAIN: A;	ELECTRON TRANSPORT ELECTRON TRANSPORT ANTIBODY ANTIBODY, FRAMEWORK
1652	1h8n	A	21	296	7.2e-50	0.12	-0.08		IGG2A KAPPA ANTIBODY CB41 (LIGHT CHAIN); CHAIN: A; IGG2A KAPPA ANTIBODY CB41 (HEAVY CHAIN); CHAIN: B; PEPTIDE 5; CHAIN: C;	COMPLEX (ANTIBODY/PEPTIDE) POLYSPECIFICITY, CROSSREACTIVITY, FAB-FRAGMENT, PEPTIDE, 2 HIV-1
1652	1hi6	B	20	258	1.2e-47	-0.01	0.01		IMMUNOGLOBULIN IGG FAB (IGG2B, KAPPA) FRAGMENT (40-50 FAB) COMPLEXED WITH IIBG 3 OUBAIN IIBG 4	
1652	1ibg	H	21	229	7e-45	0.25	0.70		LYSOZYME BINDING IGG KAPPA CHAIN; CHAIN: L; IGG1 FAB CHAIN H; CHAIN: H; LYSOZYME C; CHAIN: Y;	PROTEIN BINDING/HYDROLASE ANTIGEN-ANTIBODY COMPLEX, HYHEL-10, ANTI-HEN EGG WHITE 2 LYSOZYME ANTIBODY
1652	1ic7	H	21	117	1.2e-61	0.71	1.00		COMPLEX(ANTIBODY-ANTIGEN) FV FRAGMENT (IGG1, KAPPA) (LIGHT AND HEAVY VARIABLE	
1652	1jhl	H	1	93	2.8e-44			63.09		

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									DOMAINS 1JHL 3 NON-COVALENTLY ASSOCIATED) OF MONOCLONAL ANTI-HEN EGG 1JHL 4 LYSOZYME ANTIBODY D1.15 COMPLEX WITH PHEASANT EGG 1JHL 5 LYSOZYME 1JHL 6	
1652	1jhl	H	21	127	6e-56	0.89	1.00		ANTIBODY A6; CHAIN: L, H; INTERFERON-GAMMA RECEPTOR ALPHA CHAIN; CHAIN: I;	COMPLEX (ANTIBODY/ANTIGEN) CYTOKINE RECEPTOR, COMPLEX (ANTIBODY/ANTIGEN), 2 TRANSMEMBRANE, GLYCOPROTEIN
1652	Imco	H	209	316	2.8e-55	0.45	0.82		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1652	Imek		7	125	5.6e-32			99.63	PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX-ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
1652	Insu	H	20	239	1.4e-49			142.58	IGG FAB (IGG1, KAPPA); INSU 4 CHAIN: L, H; INSU 5 STAPHYLOCOCCAL NUCLEASE; INSU 9 CHAIN: S; INSU 10	COMPLEX (IMMUNOGLOBULIN/HYDROLASE) N10 FAB IMMUNOGLOBULIN; INSU 7 STAPHYLOCOCCAL NUCLEASE; RIBONUCLEASE, INSU 11 IMMUNOGLOBULIN, STAPHYLOCOCCAL NUCLEASE INSU 25
1652	Iosp	H	20	240	2.8e-45			136.38	FAB 184.1; CHAIN: L, H; OUTER SURFACE PROTEIN A; CHAIN: O;	COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN) OSPA; COMPLEX (IMMUNOGLOBULIN/LIPOPROTEI

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1652	1qkz	H	21	172	2.8e-34	0.36	0.53		ANTIBODY; CHAIN: H, L; PROTEIN G-PRIME; CHAIN: A; MAJOR OUTER MEMBRANE PROTEIN P1.16; CHAIN: P;	N), OUTER SURFACE 2 PROTEIN A COMPLEXED WITH FAB184.1, BORRELLIA BURGDORFERI 3 STRAIN B31
1652	1sm3	H	20	172	2.8e-34	0.45	0.47		SM3 ANTIBODY; CHAIN: L, H; PEPTIDE EPTOPE; CHAIN: P;	IMMUNE SYSTEM FAB, POR A, NEISSERIA MENINGITIDIS, PORIN
1652	2trx	A	11	126	2.8e-33			80.58	ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2 2TRXA 3	COMPLEX (ANTIBODY/PEPTIDE EPTOPE) ANTIBODY, PEPTIDE ANTIGEN, ANTITUMOR ANTIBODY, 2 COMPLEX (ANTIBODY/PEPTIDE EPTOPE)
1652	32c2	B	207	316	5.6e-55	0.49	0.99		IGG1 ANTIBODY 32C2; CHAIN: A; IGG1 ANTIBODY 32C2; CHAIN: B;	IMMUNE SYSTEM FAB, ANTIBODY, AROMATASE, P450
1652	7fab	H	1	104	2.8e-46			78.15	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
1652	7fab	H	207	316	4.2e-54	0.24	0.99		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
1652	7fab	H	20	224	1.3e-52			158.83	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	

972

Table 5

SEQ ID NO.	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1658	1a0r	P	389	562	9.6e-30	0.12	-0.17		TRANSDUCIN; CHAIN: B, G, PHOSDUCIN; CHAIN: P;	COMPLEX (TRANSDUCER/TRANSDUCTION) GT BETA-GAMMA; MEKA, PP33 PHOSDUCIN, TRANSDUCIN, BETA-GAMMA, SIGNAL TRANSDUCTION, 2 REGULATION, PHOSPHORYLATION, G PROTEINS, THIOREDOXIN, 3 VISION, MEKA, COMPLEX (TRANSDUCER/TRANSDUCTION), 4 POST-TRANSLATIONAL MODIFICATION, FARNESYL, FARNESYLATION HEADER HETNAM
1658	1a8l		324	560	5.6e-15	0.14	0.25		PROTEIN DISULFIDE OXIDOREDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, PDI, THIOREDOXIN FOLD
1658	1a8l		338	570	1.1e-41	0.31	0.78		PROTEIN DISULFIDE OXIDOREDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, PDI, THIOREDOXIN FOLD
1658	1a8l		463	690	7e-20	0.20	0.10		PROTEIN DISULFIDE OXIDOREDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, PDI, THIOREDOXIN FOLD
1658	1a8y		319	689	2.8e-19	0.24	0.09		CALSEQUESTRIN; CHAIN: NULL	CALCIUM-BINDING PROTEIN CALSEQUESTRIN, CALCIUM-BINDING PROTEIN, SARCOPLASMIC 2 RETICULUM, RABBIT SKELETAL MUSCLE
1658	1a8y		321	630	1.2e-39	0.08	0.90		CALSEQUESTRIN; CHAIN: NULL	CALCIUM-BINDING PROTEIN CALSEQUESTRIN, CALCIUM-BINDING PROTEIN, SARCOPLASMIC 2 RETICULUM, RABBIT SKELETAL MUSCLE
1658	1eqg	A	462	554	1.4e-22	0.34	0.99		THIOREDOXIN; CHAIN: A; REF-1 PEPTIDE;	COMPLEX (ELECTRON TRANSPORT/PEPTIDE) COMPLEX,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1658	1cgg	A	575	688	9.8e-20	0.21	-0.09		CHAIN: B; THIOREDOXIN; CHAIN: A; REF-1 PEPTIDE; CHAIN: B;	ELECTRON TRANSPORT/PEPTIDE COMPLEX (ELECTRON TRANSPORT/PEPTIDE) COMPLEX, ELECTRON TRANSPORT/PEPTIDE
1658	1dby	A	13	119	8.4e-26			66.72	CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
1658	1dby	A	326	431	6e-29	0.77	1.00		CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
1658	1dby	A	330	432	2.8e-25	0.93	1.00		CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
1658	1dby	A	461	570	4.8e-30	0.61	1.00		CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
1658	1dby	A	465	568	8.4e-26	0.60	1.00		CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
1658	1ev		462	554	5.6e-25	0.39	0.96		THIOREDOXIN; CHAIN: NULL;	OXIDOREDUCTASE DIMER, THIOREDOXIN, X-RAY CRYSTALLOGRAPHY, OXIDOREDUCTASE
1658	1f9m	A	7	119	2.8e-18			56.09	THIOREDOXIN F; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
1658	1faa	A	1	119	2.8e-18			55.11	THIOREDOXIN F; CHAIN: A;	ELECTRON TRANSPORT ELECTRON TRANSPORT
1658	1fb6	A	14	118	1.4e-27			67.56	THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1658	1fb6	A	327	431	1.4e-26	0.85	1.00		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
1658	1fb6	A	327	431	2.4e-28	0.85	1.00		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
1658	1fb6	A	462	563	1.1e-29	0.22	1.00		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
1658	1fb6	A	462	568	7e-27	0.17	1.00		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
1658	1fb6	A	578	689	1.3e-23	0.47	-0.05		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
1658	1mek		321	435	2.8e-26	0.50	0.83		PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX- ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
1658	1mek		322	431	2.4e-26	0.82	1.00		PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX- ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
1658	1mek		455	573	2.4e-28			99.73	PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX- ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
1658	1mek		457	562	2.4e-28	1.03	1.00		PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX- ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
1658	1mek		587	687	1.4e-18	0.22	0.24		PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX- ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
1658	1mek		6	122	8.4e-27			93.85	PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX- ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM

975

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1658	1quw	A	13	120	1.4e-24			61.41	THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA OPEN-TWISTED PROTEIN, THIOL-DISULFIDE
1658	1quw	A	327	432	9.6e-27	0.91	1.00		THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA OPEN-TWISTED PROTEIN, THIOL-DISULFIDE
1658	1quw	A	465	558	4.2e-26	0.44	1.00		THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA OPEN-TWISTED PROTEIN, THIOL-DISULFIDE
1658	1quw	A	594	689	1.1e-22	0.28	-0.12		THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA OPEN-TWISTED PROTEIN, THIOL-DISULFIDE
1658	1thx		11	119	1.1e-22			64.88	THIOREDOXIN; 1THX 5 CHAIN: NULL; 1THX 6	ELECTRON TRANSPORT THIOREDOXIN 2; 1THX 7 OXIDO-REDUCTASE 1THX 16
1658	1thx		324	431	3.6e-29	0.84	1.00		THIOREDOXIN; 1THX 5 CHAIN: NULL; 1THX 6	ELECTRON TRANSPORT THIOREDOXIN 2; 1THX 7 OXIDO-REDUCTASE 1THX 16
1658	1thx		575	689	8.4e-19	0.37	0.36		THIOREDOXIN; 1THX 5 CHAIN: NULL; 1THX 6	ELECTRON TRANSPORT THIOREDOXIN 2; 1THX 7 OXIDO-REDUCTASE 1THX 16
1658	2trx	A	11	116	9.8e-28			75.24	ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2	
1658	2trx	A	324	431	4.8e-30	0.94	1.00		ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2	
1658	2trx	A	324	431	8.4e-27	1.06	1.00		ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2	
1658	2trx	A	461	563	7.2e-30	0.43	1.00		ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2	
1658	2trx	A	462	565	8.4e-27	0.37	1.00		ELECTRON TRANSPORT	

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									THIOREDOXIN 2TRXA 2 2TRXA 3	
1658	2trx	A	575	690	2.8e-24	0.45	0.15		ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2 2TRXA 3	
1662	1c3p	A	1134	1434	1.4e-69	0.31	0.98		HDLP (HISTONE DEACETYLASE-LIKE PROTEIN); CHAIN: A;	LYASE ALPHA/BETA FOLD, LYASE
1666	1en2	A	31	113	1.4e-06	0.17	-0.02		AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/ CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTGEN, SACCHARIDE BINDING
1666	2rel		11	67	8.4e-23			111.70	R-ELAFIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, R- ELAFIN, ELASTASE INHIBITOR
1666	2rel		61	117	4.8e-24	-0.64	1.00		R-ELAFIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, R- ELAFIN, ELASTASE INHIBITOR
1666	2rel		61	117	4.8e-24			111.59	R-ELAFIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, R- ELAFIN, ELASTASE INHIBITOR
1666	2rel		62	116	1.4e-11	-0.45	1.00		R-ELAFIN; CHAIN: NULL;	SERINE PROTEASE INHIBITOR SERINE PROTEASE INHIBITOR, R- ELAFIN, ELASTASE INHIBITOR
1672	1b1h	A	132	426	2.8e-24	0.23	0.23		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1672	1b1h	A	20	410	4.2e-46	0.11	-0.11		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1672	1b1h	A	30	422	4.2e-46			96.50	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1672	1edy		140	310	3.6e-20	0.26	0.24		T-CELL SURFACE	T-CELL SURFACE



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									GLYCOPROTEIN CD4 <sub>5</sub> CHAIN: NULL;	GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
1672	1cs6	A	20	422	1.4e-59	0.30	0.78		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1672	1cs6	A	25	405	1.4e-59			104.76	AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1672	1cvs	C	126	300	8.4e-29	0.01	0.35		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1672	1cvs	C	224	421	2.8e-41	0.35	0.01		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1672	1cvs	C	328	441	1.4e-15	0.05	-0.19		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1672	1cvs	D	124	310	9.8e-30			65.66	FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1672	1cvs	D	126	300	9.8e-30	0.10	0.96		CHAIN: C, D; FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1672	1cvs	D	224	421	1.4e-37	0.42	0.69		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1672	1cvs	D	224	421	1.4e-37	0.42	0.69		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1672	1epf	A	132	310	3.6e-23	0.44	0.34		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1672	1epf	A	137	294	2.8e-19	0.22	0.86		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1672	1epf	A	21	202	7e-20	0.18	-0.17		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1672	1epf	A	235	421	1.4e-17	0.30	-0.03		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1672	1ev2	E	130	300	1.4e-24	-0.00	0.45		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2, FGFR2, IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1672	1ev2	E	226	421	1.1e-34	0.32	-0.06		F, G, H; FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	DOMAINS, B-TREFOIL FOLD GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1672	1ev2	E	328	438	7e-15	0.00	-0.20		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1672	1ev2	G	125	311	2.4e-23	0.25	0.96		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1672	1ev2	G	226	425	1.1e-36	0.18	0.01		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1672	1evt	C	122	311	1.3e-29			72.46	FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1672	1evt	C	126	300	1.3e-29	0.00	1.00		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: C, D;	SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1672	1evt	C	224	421	1.4e-35	0.54	-0.11		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1672	1f2q	A	138	293	4.2e-17	0.13	0.47		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1672	1f2q	A	20	223	1.4e-17	0.22	0.94		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1672	1f6a	A	130	293	5.6e-17	0.25	0.99		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1672	1f6a	A	16	223	1.1e-17	0.05	0.09		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1672	1f97	A	238	420	1.4e-22	0.25	0.12		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD

981

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1672	1f97	A	27	215	1.4e-34	0.32	-0.11		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1672	1fcg	A	128	293	4.2e-15	0.20	0.98		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1672	1fcg	A	22	219	1.2e-18	-0.05	0.01		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1672	1fcg	A	247	421	1.2e-15	-0.25	0.03		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1672	1fhg	A	227	300	2.8e-13	0.16	0.16		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
1672	1fhl	A	22	219	3.6e-18	-0.06	0.03		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1672	1fhl	A	240	421	2.4e-15	0.25	-0.08		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1672	1g0x	A	130	297	4.8e-22	0.19	0.93		LEUCOCYTE IMMUNOGLOBULIN-LIKE RECEPTOR-1; CHAIN: A;	IMMUNE SYSTEM LEUCOCYTE INHIBITORY RECEPTOR-1; LEUKOCYTE IMMUNOGLOBULIN FOLD, 3-10 HELIX
1672	1iam		139	310	1.2e-14	0.00	0.07		INTERCELLULAR ADHESION MOLECULE-1; CHAIN: NULL;	RHINOVIRUS RECEPTOR ICAM-1, CD54; RHINOVIRUS RECEPTOR, CELL ADHESION, INTEGRIN

Table 5

SEQ ID No:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1672	1ie5	A	215	311	4.8e-13	-0.22	0.25		NEURAL CELL ADHESION MOLECULE; CHAIN: A;	LIGAND, 2 GLYCOPROTEIN, LFA-1 LIGAND, IMMUNOGLOBULIN FOLD, 3 TRANSMEMBRANE CELL ADHESION N-CAM; INTERMEDIATE IMMUNOGLOBULIN FOLD
1672	1igt	B	5	437	1.4e-10			79.51	IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1672	1iil	G	226	425	2.8e-41	0.38	0.21		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
1672	1iil	G	31	223	2.8e-26	0.11	-0.14		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
1672	1itb	B	134	428	4.2e-29			69.34	INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1672	1itb	B	151	419	4.2e-29	0.02	0.17		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1672	1ib	B	24	310	1.2e-24	0.16	0.17		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR) COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1672	1nkr		131	313	1.2e-17			70.67	P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
1672	1nkr		133	297	1.2e-17	0.05	0.82		P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
1672	1nkr		23	219	1.2e-16	0.21	0.00		P58-CL42 KIR; CHAIN: NULL;	INHIBITORY RECEPTOR KILLER CELL INHIBITORY RECEPTOR; INHIBITORY RECEPTOR, NATURAL KILLER CELLS, IMMUNOLOGICAL 2 RECEPTORS, IMMUNOGLOBULIN FOLD
1672	1vca	A	242	353	2.4e-13	-0.04	0.01		HUMAN VASCULAR CELL ADHESION MOLECULE-1; 1VCA 4 CHAIN: A, B; 1VCA 5	CELL ADHESION PROTEIN VCAM-D1,2; 1VCA 6 IMMUNOGLOBULIN SUPERFAMILY, INTEGRIN-BINDING 1VCA 15
1672	1wio	A	27	410	1.2e-27			92.72	T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1672	1wio	A	27	417	1.2e-27	0.04	-0.14		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	LIPOPROTEIN, POLYMORPHISM GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
1672	2di	A	131	316	1.2e-22			66.58	MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
1672	2di	A	133	312	1.2e-22	0.20	0.82		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
1672	2di	A	23	219	6e-19	0.14	-0.09		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR; KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
1672	2fcb	A	128	293	1.4e-15	0.29	0.68		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1672	2fcb	A	247	421	2.4e-15	0.07	0.04		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1672	8fab	A	25	211	5.6e-17	0.21	-0.14		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
1673	1a25	A	507	632	7e-27	0.32	0.47		PROTEIN KINASE C	CALCIUM-BINDING PROTEIN



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(BETA); CHAIN: A, B;	CALB; CALCIUM <sup>++</sup> /PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN
1673	1byn	A	502	630	1.4e-27	0.12	0.95		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1673	1dix	A	491	619	5.6e-20	0.05	-0.15		PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C, CHAIN: A, B;	LIPID DEGRADATION PLC-D1; PHOSPHORIC DIESTER HYDROLASE, HYDROLASE, LIPID DEGRADATION, 2 TRANSDUCER, CALCIUM-BINDING, PHOSPHOLIPASE C, 3 PHOSPHOINOSITIDE-SPECIFIC
1673	1dqy	A	500	646	2.8e-29	0.21	0.90		SYNAPTOTAGMIN III; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS BETA SANDWICH, CALCIUM ION, C2 DOMAIN
1673	1dsy	A	507	632	1.3e-26	-0.05	0.42		PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	TRANSFERASE CALCIUM <sup>++</sup> , PHOSPHOLIPID BINDING PROTEIN, CALCIUM-BINDING 2 PROTEIN, PHOSPHATIDYL SERINE, PROTEIN KINASE C
1673	1rsy		497	627	2.8e-27	0.25	0.93		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
1673	3rpb	A	522	639	7e-25	0.08	0.04		RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1678	1ao7	E	22	154	2.8e-42	0.05	0.88		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS I MHC, T-CELL RECEPTOR, VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
1678	1bd2	E	22	154	9.8e-44	-0.05	0.48		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)
1678	1bec		23	154	1.4e-41	-0.03	0.51		14.3.D T CELL ANTIGEN RECEPTOR; 1BEC 5 CHAIN: NULL; 1BEC 6	RECEPTOR T CELL RECEPTOR 1BEC 14
1678	1bj1	J	20	158	5.6e-43	0.00	-0.19		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
1678	1d9k	B	22	136	1.4e-40			53.60	T-CELL RECEPTOR D10 (ALPHA CHAIN); CHAIN: A, E; T-CELL RECEPTOR D10 (BETA CHAIN); CHAIN: B, F; MHC I-AK A CHAIN (ALPHA CHAIN); CHAIN: C, G; MHC I-AK B CHAIN (BETA CHAIN);	IMMUNE SYSTEM MHC I-AK; MHC I-AK; T-CELL RECEPTOR, MHC CLASS II, D10, I-AK

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1678	1dee	A	20	158	8.4e-43	0.17	-0.19		CHAIN: D, H; CONALBUMIN PEPTIDE; CHAIN: P, Q; IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
1678	1fgy	L	20	153	8.4e-42	0.20	-0.14		IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 IFGV 3 ANTIBODY 'H52' (HUH52-AA FV) IFGV 4	
1684	1edo	A	3	261	4.5e-73			58.51	BETA-KETO ACYL CARRIER PROTEIN REDUCTASE; CHAIN: A;	OXIDOREDUCTASE NUCLEOTIDE FOLD, ROSSMANN FOLD
1684	1edo	A	5	233	4.5e-73	0.52	-0.20		BETA-KETO ACYL CARRIER PROTEIN REDUCTASE; CHAIN: A;	OXIDOREDUCTASE NUCLEOTIDE FOLD, ROSSMANN FOLD
1684	1fds		2	241	6e-73	0.75	-0.20		17-BETA- HYDROXYSTEROID- DEHYDROGENASE; CHAIN: NULL;	DEHYDROGENASE DEHYDROGENASE, 17-BETA- HYDROXYSTEROID
1684	1fds		2	242	6e-73			304.79	17-BETA- HYDROXYSTEROID- DEHYDROGENASE; CHAIN: NULL;	DEHYDROGENASE DEHYDROGENASE, 17-BETA- HYDROXYSTEROID
1685	1erg		132	212	0.00053	0.03	-0.20		COMPLEMENT FACTOR HUMAN COMPLEMENT REGULATORY PROTEIN CD59 (EXTRACELLULAR IERG 3 REGION,	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1685	1erg		24	107	0.0018	0.08	-0.20		RESIDUES 1 - 70) (NMR, RESTRAINED MINIMIZED 1ERG 4 AVERAGE STRUCTURE) 1ERG 5	
1686	1b6c	B	1	238	3e-75			127.53	COMPLEMENT FACTOR HUMAN COMPLEMENT REGULATORY PROTEIN CD59 (EXTRACELLULAR 1ERG 3 REGION, RESIDUES 1 - 70) (NMR, RESTRAINED MINIMIZED 1ERG 4 AVERAGE STRUCTURE) 1ERG 5	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
1686	1b6c	B	523	908	1.5e-97			202.68	FK506-BINDING PROTEIN; CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
1686	1b6c	B	526	907	1.5e-97	0.37	-0.20		FK506-BINDING PROTEIN; CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	COMPLEX (ISOMERASE/PROTEIN KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1686	1blx	A	559	852	3e-28	0.18	-0.20		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	KINASE COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
1686	1byg	A	551	772	9e-32	0.08	-0.20		C-TERMINAL SRC KINASE; CHAIN: A;	TRANSFERASE CSK; PROTEIN KINASE, C-TERMINAL SRC KINASE, PHOSPHORYLATION, 2 STAUROSPORINE, TRANSFERASE
1686	1e1x	A	552	770	9e-30	0.23	-0.20		CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: A;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, CELL DIVISION, 2 MITOSIS, INHIBITION
1686	1f3m	C	557	772	1.3e-28	0.24	-0.20		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A; B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
1686	1fgk	A	553	770	1.8e-32	0.15	-0.20		FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
1686	1fgk	A	575	843	3e-32	0.17	-0.20		FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR,

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1686	1fmk		553	836	3e-38	0.06	-0.20		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE PHOSPHOTRANSFERASE C-SRC, P60-SRC, SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
1686	1hcl		552	770	1.5e-30	0.02	-0.20		HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
1686	1i44	A	552	843	9e-30	0.16	-0.20		INSULIN RECEPTOR; CHAIN: A;	TRANSFERASE IR- PROTEIN TYROSINE KINASE, PHOSPHOTRANSFERASE
1686	1i44	A	553	770	2.4e-30	0.14	-0.20		INSULIN RECEPTOR; CHAIN: A;	TRANSFERASE IR- PROTEIN TYROSINE KINASE, PHOSPHOTRANSFERASE
1686	1iep	A	553	842	3e-38	0.08	-0.20		PROTO-ONCOGENE TYROSINE-PROTEIN KINASE ABL; CHAIN: A, B;	TRANSFERASE P150, C-ABL; KINASE, KINASE INHIBITOR, STI-571, ACTIVATION LOOP
1686	1i13	A	552	843	3e-31	0.17	-0.20		INSULIN RECEPTOR; CHAIN: A; PEPTIDE SUBSTRATE; CHAIN: B;	COMPLEX (TRANSFERASE/SUBSTRATE) TYROSINE KINASE, SIGNAL TRANSDUCTION, PHOSPHOTRANSFERASE, 2 COMPLEX (KINASE/PEPTIDE SUBSTRATE/ATP ANALOG), ENZYME, 3 COMPLEX (TRANSFERASE/SUBSTRATE)
1686	1i13	A	553	770	5.9e-33	0.40	-0.20		INSULIN RECEPTOR; CHAIN: A; PEPTIDE	COMPLEX (TRANSFERASE/SUBSTRATE)

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									SUBSTRATE; CHAIN: B;	TYROSINE KINASE, SIGNAL TRANSDUCTION, PHOSPHOTRANSFERASE, 2 COMPLEX (KINASE/PEPTIDE SUBSTRATE/ATP ANALOG), ENZYME, 3 COMPLEX (TRANSFERASE/SUBSTRATE)
1686	1phk		552	772	1.1e-30	0.27	-0.20		PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
1686	1qcf	A	553	836	3e-33	0.04	-0.20		HAEMATOPPOETIC CELL KINASE (HCK); CHAIN: A;	TYROSINE KINASE TYROSINE KINASE-INHIBITOR COMPLEX, DOWN-REGULATED KINASE, 2 ORDERED ACTIVATION LOOP
1686	1vr2	A	553	770	4.7e-32	0.11	-0.20		VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR CHAIN: A;	TRANSFERASE KDR; TYROSINE KINASE
1688	1fmk		149	233	2.9e-29	0.04	-0.20		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
1688	1fmk		149	233	4.5e-26	0.04	-0.20		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
1688	1fyn	A	146	207	2.4e-18			82.54	PHOSPHOTRANSFERASE	TRANSFERASE PROTO-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FYN; CHAIN: A; 3BP-2; CHAIN: B;	ONCOGENE TYROSINE KINASE; PROTO-ONCOGENE, TRANSFERASE, TYROSINE-PROTEIN KINASE, 2 PHOSPHORYLATION, ATP-BINDING, MYRISTYLATION, SH3 DOMAIN, 3 COMPLEX (PHOSPHOTRANSFERASE/PEPTID E)
1688	1fyn	A	148	207	2.4e-18	0.53	-0.20		PHOSPHOTRANSFERASE FYN; CHAIN: A; 3BP-2; CHAIN: B;	TRANSFERASE PROTO-ONCOGENE TYROSINE KINASE; PROTO-ONCOGENE, TRANSFERASE, TYROSINE-PROTEIN KINASE, 2 PHOSPHORYLATION, ATP-BINDING, MYRISTYLATION, SH3 DOMAIN, 3 COMPLEX (PHOSPHOTRANSFERASE/PEPTID E)
1688	1fyn	A	7	68	1.5e-21			86.38	PHOSPHOTRANSFERASE FYN; CHAIN: A; 3BP-2; CHAIN: B;	TRANSFERASE PROTO-ONCOGENE TYROSINE KINASE; PROTO-ONCOGENE, TRANSFERASE, TYROSINE-PROTEIN KINASE, 2 PHOSPHORYLATION, ATP-BINDING, MYRISTYLATION, SH3 DOMAIN, 3 COMPLEX (PHOSPHOTRANSFERASE/PEPTID E)
1688	1g2b	A	137	188	1.8e-14	0.22	-0.20		SPECTRIN ALPHA CHAIN; CHAIN: A;	METAL BINDING PROTEIN CAPPING PROTEIN, CALCIUM-BINDING, DUPLICATION, REPEAT, SH3 2 DOMAIN, CYTOSKELETON
1688	1g83	A	150	233	1.2e-16	0.08	-0.20		PROTO-ONCOGENE TYROSINE-PROTEIN	TRANSFERASE P59-FYN; BETA BARREL, ANTIPARALLEL BETA



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									KINASE FYN; CHAIN: A, B;	SHEET, ALPHA HELIX, 3-10 2 HELIX
1688	1g83	A	150	233	3e-29			73.74	PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FYN; CHAIN: A, B;	TRANSFERASE P59-FYN; BETA BARREL, ANTIPARALLEL BETA SHEET, ALPHA HELIX, 3-10 2 HELIX
1688	1g83	A	152	233	3e-29	0.47	-0.20		PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FYN; CHAIN: A, B;	TRANSFERASE P59-FYN; BETA BARREL, ANTIPARALLEL BETA SHEET, ALPHA HELIX, 3-10 2 HELIX
1688	1gri	A	6	208	3e-05			50.90	GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B, 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
1688	1lii	A	131	213	1.2e-15	0.07	-0.20		MELANOMA DERIVED GROWTH REGULATORY PROTEIN; CHAIN: A, B;	HORMONE/GROWTH FACTOR MELANOMA INHIBITORY ACTIVITY; SH3 SUBDOMAIN
1688	1qcf	A	148	233	2.4e-32	0.01	-0.20		HAEMATOPOIETIC CELL KINASE (HCK); CHAIN: A;	TYROSINE KINASE TYROSINE KINASE-INHIBITOR COMPLEX, DOWN-REGULATED KINASE, 2 ORDERED ACTIVATION LOOP
1688	1shf	A	10	68	3e-21			83.69	PHOSPHOTRANSFERASE FYN PROTO-ONCOGENE TYROSINE KINASE (E.C.2.7.1.112) ISHF 3 (SH3 DOMAIN) ISHF 4	
1688	1shf	A	149	207	1.2e-18	0.28	-0.20		PHOSPHOTRANSFERASE FYN PROTO-ONCOGENE TYROSINE KINASE (E.C.2.7.1.112) ISHF 3 (SH3 DOMAIN) ISHF 4	
1688	1shf	A	149	207	1.2e-18			79.90	PHOSPHOTRANSFERASE FYN PROTO-ONCOGENE TYROSINE KINASE (E.C.2.7.1.112) ISHF 3	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1688	1tud		137	188	1.2e-14	0.37	-0.20		(SH3 DOMAIN) 1SHF 4 ALPHA-SPECTRIN; CHAIN: NULL;	CYTOSKELETON CAPPING PROTEIN, CALCIUM-BINDING, DUPLICATION, REPEAT, 2 SH3 DOMAIN, CYTOSKELETON
1693	1a4y	A	506	751	5.9e-30	0.19	-0.20		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE- RICH 3 REPEATS
1693	1a9n	A	492	587	1.2e-14	0.15	-0.20		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1a9n	A	511	671	1.2e-22	0.07	-0.20		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1a9n	A	619	749	1.2e-16	0.32	-0.20		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1a9n	A	669	760	3e-09	0.17	-0.20		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1a9n	C	494	624	3.5e-17	0.32	-0.20		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1a9n	C	511	702	2.4e-19	0.10	-0.20		U2 RNA HAIRPIN IV;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	PROTEIN/RNA COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1a9n	C	619	749	5.9e-17	0.23	-0.20		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1a9n	C	664	760	3.5e-09	0.21	-0.20		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1693	1d0b	A	481	640	3e-14	0.22	-0.20		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1693	1d0b	A	490	642	5.9e-18	0.15	-0.20		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1693	1d0b	A	504	696	5.3e-20	0.02	-0.20		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1693	1d0b	A	514	747	1.8e-23	0.10	-0.20		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1693	1d0b	A	525	673	4.5e-16	0.41	-0.20		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1693	1ds9	A	613	749	1.2e-13	0.13	-0.20		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1693	1fs2	A	527	750	5.9e-12	0.20	-0.20		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1,

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1693	1jvt		396	433	0.0012	0.04	-0.20		HUMAN T-CELL LEUKEMIA VIRUS TYPE II MATRIX CHAIN: NULL;	SKP2, F-BOX, LRKS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1693	1yrg	A	503	750	4.1e-28	0.15	-0.20		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1693	2bnh		504	749	3.5e-32	0.34	-0.20		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1694	1cka	A	124	170	0.001	0.23	-0.20		COMPLEX (ONCOGENE PROTEIN/PEPTIDE) C-CRK (N-TERMINAL SH3 DOMAIN) (C-CRKSH3-N) COMPLEXED WITH ICKA 3 C3G PEPTIDE (PRO-PRO-PRO-ALA-LEU-PRO-PRO-LYS-LYS-ARG) ICKA 4	
1694	1cka	A	19	73	1e-12	0.12	-0.20		COMPLEX (ONCOGENE PROTEIN/PEPTIDE) C-CRK (N-TERMINAL SH3	

997

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1694	1efh	A	21	70	6e-13	0.14	-0.20		DOMAIN (C-CRKS13-N) COMPLEXED WITH 1CKA 3 C3G PEPTIDE (PRO-PRO-PRO-ALA-LEU-PRO-PRO-LYS-LYS-ARG) 1CKA 4	
1694	1fyn	A	17	70	4.5e-14	0.36	-0.20		FYN TYROSINE KINASE; CHAIN: A, C; HIV-1 NEF PROTEIN; CHAIN: B, D;	COMPLEX (SH3 DOMAIN/VIRAL ENHANCER) SRC-HOMOLOGY 3 DOMAIN; COMPLEX (SH3 DOMAIN/VIRAL ENHANCER), PROTO-ONCOGENE, 2 TRANSFERASE, TYROSINE-PROTEIN KINASE, PHOSPHORYLATION, 3 AIDS, MYRISTYLATION, GTP-BINDING, ATP-BINDING, SH3 DOMAIN, 4 SH2 DOMAIN, PPII HELIX, PXXP MOTIF
1694	1fyn	A	17	70	4.5e-14	0.36	-0.20		PHOSPHOTRANSFERASE FYN; CHAIN: A; 3BP-2; CHAIN: B;	TRANSFERASE PROTO-ONCOGENE TYROSINE KINASE; PROTO-ONCOGENE, TRANSFERASE, TYROSINE-PROTEIN KINASE, 2 PHOSPHORYLATION, ATP-BINDING, MYRISTYLATION, SH3 DOMAIN, 3 COMPLEX (PHOSPHOTRANSFERASE/PEPTID E)
1694	1g83	A	21	80	6e-14	0.20	-0.20		PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FYN; CHAIN: A, B;	TRANSFERASE P59-FYN; BETA BARREL, ANTIPARALLEL BETA SHEET, ALPHA HELIX, 3-10 2 HELIX
1694	1gbq	A	21	74	4.5e-12	0.33	-0.20		GRB2; CHAIN: A; SOS-1; CHAIN: B;	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) COMPLEX (SIGNAL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1694	1gtr	A	123	176	0.001	0.31	-0.20		SIGNAL TRANSDUCTION PROTEIN GROWTH FACTOR RECEPTOR- BOUND PROTEIN 2 (GRB2, N-TERMINAL IGBR 3 SH3 DOMAIN) COMPLEXED WITH SOS- A PEPTIDE IGBR 4 (NMR, 29 STRUCTURES) IGBR 5	TRANSDUCTION/PEPTIDE, SH3 DOMAIN
1694	1gtr	A	21	82	1.3e-12	0.28	-0.20		SIGNAL TRANSDUCTION PROTEIN GROWTH FACTOR RECEPTOR- BOUND PROTEIN 2 (GRB2, N-TERMINAL IGBR 3 SH3 DOMAIN) COMPLEXED WITH SOS- A PEPTIDE IGBR 4 (NMR, 29 STRUCTURES) IGBR 5	
1694	1gcq	A	125	173	1.5e-07	0.40	-0.20		GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; CHAIN: A, B; VAV PROTO- ONCOGENE; CHAIN: C;	SIGNALING PROTEIN/SIGNALING PROTEIN SH3 DOMAIN, PROTEIN- PROTEIN COMPLEX, GRB2, VAV
1694	1gcq	A	20	70	1.5e-14	0.24	-0.20		GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; CHAIN: A, B; VAV PROTO- ONCOGENE; CHAIN: C;	SIGNALING PROTEIN/SIGNALING PROTEIN SH3 DOMAIN, PROTEIN- PROTEIN COMPLEX, GRB2, VAV
1694	1gcq	B	125	173	1.5e-07	0.05	-0.20		GROWTH FACTOR	SIGNALING PROTEIN/SIGNALING

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									RECEPTOR-BOUND PROTEIN 2; CHAIN: A, B; VAV PROTO-ONCOGENE; CHAIN: C;	PROTEIN SH3 DOMAIN, PROTEIN-PROTEIN COMPLEX, GRB2, VAV
1694	1gcq	B	20	70	1.5e-14	0.13	-0.20		GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; CHAIN: A, B; VAV PROTO-ONCOGENE; CHAIN: C;	SIGNALING PROTEIN/SIGNALING PROTEIN SH3 DOMAIN, PROTEIN-PROTEIN COMPLEX, GRB2, VAV
1694	1gfc		17	70	1.3e-14	0.30	-0.20		ADAPTOR PROTEIN CONTAINING SH2 AND SH3 GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2) 1GFC 3 (C-TERMINAL SH3 DOMAIN) (NMR, MINIMIZED MEAN STRUCTURE) 1GFC 4	
1694	1gri	A	15	70	4.5e-15	0.16	-0.20		GROWTH FACTOR BOUND PROTEIN 2: 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
1694	1hsq		21	82	4.5e-12	0.21	-0.20		PHOSPHORIC DIESTER HYDROLASE PHOSPHOLIPASE C-GAMMA (SH3 DOMAIN) (E.C.3.1.4.11) 1HSQ 3 (NMR, MINIMIZED MEAN STRUCTURE) 1HSQ 4	
1694	1pwt		17	70	1.5e-12	0.55	-0.20		ALPHA SPECTRIN; CHAIN: NULL;	CIRCULAR PERMUTANT PWT; CIRCULAR PERMUTANT, SH3 DOMAIN, CYTOSKELETON
1694	1qcf	A	21	80	1.5e-11	0.28	-0.20		HAEMATOPHOETIC CELL KINASE (HCK); CHAIN: A;	TYROSINE KINASE TYROSINE KINASE-INHIBITOR COMPLEX, DOWN-REGULATED KINASE, 2

1000

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1694	1sem	A	19	73	1.5e-15	0.12	-0.20		SEM-5; 1SEM 3 CHAIN: A, B; 1SEM 5 10-RESIDUE PROLINE-RICH PEPTIDE FROM MSOS 1SEM 8 CHAIN: C, D 1SEM 10	ORDERED ACTIVATION LOOP SIGNAL TRANSDUCTION PROTEIN SRC-HOMOLOGY 3 (SH3) DOMAIN, PEPTIDE-BINDING PROTEIN, 1SEM 18 2 GUANINE NUCLEOTIDE EXCHANGE FACTOR 1SEM 19
1694	1shf	A	18	70	6e-14	0.36	-0.20		PHOSPHOTRANSFERASE FYN PROTO-ONCOGENE TYROSINE KINASE (E.C.2.7.1.112) 1SHF 3 (SH3 DOMAIN) 1SHF 4	
1694	1yrs	B	23	74	6e-12	0.34	-0.20		P53; CHAIN: A; 53BP2; CHAIN: B;	COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS) P53BP2; ANKYRIN REPEATS, SH3, P53, TUMOR SUPPRESSOR, MULTIGENE 2 FAMILY, NUCLEAR PROTEIN, PHOSPHORYLATION, DISEASE MUTATION, 3 POLYMORPHISM, COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS)
1694	2abl		13	80	4.5e-10	0.44	-0.20		ABL TYROSINE KINASE; CHAIN: NULL;	TRANSFERASE TRANSFERASE, TYROSINE KINASE, SH3, SH2, ONCOPROTEIN
1695	1jvr		180	216	0.0024	0.05	-0.20		HUMAN T-CELL LEUKEMIA VIRUS TYPE II MATRIX CHAIN: NULL;	MATRIX PROTEIN HTLV-II MA, MA; HUMAN T-CELL LEUKEMIA VIRUS TYPE II MATRIX PROTEIN, HTLV-II 2 MATRIX PROTEIN, HTLV-II MA, RETROVIRAL MATRIX PROTEIN, P17
1706	1a4y	A	647	870	5.9e-24	0.61	0.62		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE),



1001

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1706	1a4y	A	669	956	4.5e-21	0.38	0.63		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1706	1a4y	A	684	869	2.4e-25	0.64	1.00		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1706	1a5c		370	490	1.2e-27	0.36	1.00		TUMOR SUPPRESSOR P16INK4A; CHAIN: NULL;	ANTI-ONCOGENE CELL CYCLE, ANTI-ONCOGENE, REPEAT, ANK REPEAT
1706	1awc	B	353	503	1.2e-39	0.59	1.00		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
1706	1awc	B	386	529	6e-36	0.23	1.00		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									1; CHAIN: B; DNA; CHAIN: D, E;	(TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
1706	1awc	B	661	808	7.5e-11	0.34	-0.18		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
1706	1awc	B	736	885	3e-24	0.13	-0.17		GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
1706	1b3u	A	135	338	0.00012	0.19	-0.05		PROTEIN PHOSPHATASE PP2A; CHAIN: A, B;	SCAFFOLD PROTEIN SCAFFOLD PROTEIN, PP2A, PHOSPHORYLATION, HEAT REPEAT
1706	1bd8		346	470	7.5e-25	0.37	1.00		P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
1706	1bd8		356	506	1.5e-31	0.21	1.00		P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
1706	1bd8		389	548	4.5e-25	0.21	0.76		P19INK4D CDK4/6	TUMOR SUPPRESSOR TUMOR

1003

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									INHIBITOR; CHAIN: NULL;	SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
1706	1bd8		419	574	1.5e-22	0.01	0.18		P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
1706	1bi7	B	370	471	5.9e-25	0.37	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; MULTIPLE TUMOR SUPPRESSOR; CHAIN: B;	COMPLEX (KINASE/ANTI-ONCOGENE) CDK6; P16INK4A, MTS1; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, MULTIPLE TUMOR SUPPRESSOR, 3 MTS1, COMPLEX (KINASE/ANTI-ONCOGENE) HEADER
1706	1blx	B	346	470	4.5e-24	0.19	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
1706	1blx	B	356	506	1.5e-30	0.29	0.99		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
1706	1blx	B	389	535	6e-24	0.43	0.87		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)

1004

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1706	1bu9	A	345	475	7.5e-27	0.38	0.88		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
1706	1bu9	A	353	508	1.2e-37	0.48	0.80		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
1706	1bu9	A	386	550	3e-30	0.25	0.60		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
1706	1bu9	A	422	574	1.2e-26	0.04	-0.08		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
1706	1d0b	A	697	870	5.9e-10	0.01	-0.03		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1706	1d9s	A	376	488	1.8e-31	0.37	0.94		CYCLIN-DEPENDENT KINASE 4 INHIBITOR B; CHAIN: A;	SIGNALING PROTEIN HELIX-TURN-HELIX, ANKYRIN REPEAT
1706	1dcq	A	375	478	3e-28	0.21	1.00		PK2-ASSOCIATED PROTEIN BETA; CHAIN: A;	METAL BINDING PROTEIN ZINC-BINDING MODULE, ANKYRIN REPEATS, METAL BINDING PROTEIN

1005

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1706	1elt	A	150	253	1.8e-05	0.11	0.06		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
1706	1elt	A	169	300	3.5e-09	-0.13	0.21		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
1706	1fch	A	166	262	1.5e-05	-0.11	0.66		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISOMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
1706	1fch	A	6	345	3.5e-14	0.07	0.15		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISOMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
1706	1fqv	A	632	869	5.9e-14	0.35	-0.03		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1706	1fs2	A	646	869	4.1e-12	0.27	0.11		SKP2; CHAIN: A, C, SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
1706	1ihb	A	345	474	3e-26	0.53	1.00		CYCLIN-DEPENDENT KINASE 6 INHIBITOR;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE

1006

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A, B;	INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
1706	1hhb	A	353	507	6e-37	0.39	1.00		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
1706	1hhb	A	386	549	1.5e-29	0.15	0.74		CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
1706	1hhg	A	171	273	5.9e-06	-0.09	0.60		CYCLOPHILIN 40; CHAIN: A;	ISOMERASE 40 KDA PEPTIDYL-PROLYL CIS-TRANS ISOMERASE; PPLASE IMMUNOPHILIN TETRAPEPTIDE
1706	1lkn	D	348	511	4.5e-42	0.30	1.00		NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFKB COMPLEX
1706	1myo		353	472	1.2e-24	0.62	0.98		MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
1706	1myo		383	489	1.8e-33	0.61	-1202.08		MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
1706	1nfi	E	347	511	7.5e-42	0.42	1.00		NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1706	Inf1	E	380	565	9e-34	0.24	0.27		NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
1706	Inf1	E	412	574	1.5e-28	-0.02	0.16		NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
1706	1yrG	A	690	870	4.1e-24	0.90	-1202.08		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMISPHERAL TWINNING, 3 MEROPEDRAL TWINNING, MEROPEDRY
1706	2bnh		620	949	1.8e-26	0.40	-1202.08		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1706	2bnh		669	967	7.5e-22	0.35	0.90		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1708	1alh	A	201	285	1.5e-30			70.74	QGSZ ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1708	1a1h	A	9	83	3e-29			65.50	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1708	1ard		203	231	9e-07	0.03	-0.20		TRANSCRIPTION REGULATION YEAST TRANSCRIPTION FACTOR ADRI (RESIDUES 102 - 130) 1ARD 3 (AMINO TERMINAL ZINC FINGER DOMAIN) (NMR, 10 STRUCTURES) 1ARD 4 (ADRI1B) 1ARD 5	
1708	1mey	C	10	83	1.5e-44			70.07	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1708	1mey	C	12	90	4.5e-39			51.68	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1708	1mey	C	202	284	3e-47			76.62	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1708	1mey	G	198	227	3e-11	0.36	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE,



1009

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1708	1ubd	C	179	284	1.5e-33			74.99	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (ZINC FINGER/DNA)
1708	1ubd	C	1	83	6e-29			59.28	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1708	2adr		11	68	1.3e-07			53.15	ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
1708	2gli	A	147	285	6e-32			65.93	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1710	1b2w	L	1252	1288	0.0047	0.36	-0.20		ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
1710	1bec		1249	1289	0.0035	0.13	-0.20		14.3.D T CELL ANTIGEN RECEPTOR; IBEC 5	RECEPTOR T CELL RECEPTOR IBEC 14

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1710	1fyt	E	1252	1288	0.0059	0.24	-0.20		CHAIN: NULL; 1BEC 6 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR CHAIN: A; HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DR-1 CHAIN: B; HEMAGGLUTININ HA1 PEPTIDE CHAIN; CHAIN: C; T-CELL RECEPTOR ALPHA CHAIN; CHAIN: D; T-CELL RECEPTOR BETA CHAIN; CHAIN: E;	IMMUNE SYSTEM HLA-DR1, DR A; HLA-DR1, DRB1 0101; TCR HA1.7 ALPHA CHAIN; TCR HA1.7 BETA CHAIN; PROTEIN-PROTEIN COMPLEX, IMMUNOGLOBULIN FOLD
1710	1hoc	A	1251	1289	0.0059	0.02	-0.20		HISTOCOMPATIBILITY ANTIGEN MURINE CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX CONSISTING IHOC 3 OF H-2D=B=, B2-MICROGLOBULIN, AND A 9-RESIDUE PEPTIDE IHOC 4	
1710	1qho	A	614	706	0.003	0.45	-0.20		ALPHA-AMYLASE; CHAIN: A;	HYDROLASE "MALTOGENIC" ALPHA AMYLASE; AMYLASE, GLYCOSIDE HYDROLASE, STARCH DEGRADATION
1711	1acp		182	260	2.9e-07			56.22	FATTY ACID SYNTHESIS PROTEIN ACYL CARRIER PROTEIN (NMR, 2 STRUCTURES) IACP 3	
1711	1acp		220	260	2.9e-07	0.21	-0.20		FATTY ACID SYNTHESIS PROTEIN ACYL CARRIER PROTEIN	

1011

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1712	1boe	A	74	116	1.2e-17	0.33	-0.20		(NMR, 2 STRUCTURES) IACP 3	
									INSULIN-LIKE GROWTH FACTOR-BINDING PROTEIN-5 CHAIN: A;	HORMONE/GROWTH FACTOR MINI-IGFBP-5, IGFBP-5, IGF, INSULIN-LIKE GROWTH FACTOR 2 BINDING PROTEIN, NMR, HORMONE/GROWTH FACTOR
1712	1icf	I	262	331	4.5e-17	0.02	-0.20		CATHERPSIN L: HEAVY CHAIN; CHAIN: A, C; CATHERPSIN L: LIGHT CHAIN; CHAIN: B, D; INVARIANT CHAIN; CHAIN: I, J;	HYDROLASE II FRAGMENT, CD74 FRAGMENT CYSTEINE PROTEINASE, CATHERPSIN, MHC CLASS II, INVARIANT 2 CHAIN, THYROGLOBULIN TYPE-1 DOMAIN
1712	1mhv		90	117	0.0016	0.03	-0.20		METALLOTHIONEIN CD-7 METALLOTHIONEIN-2 (ALPHA DOMAIN) (NMR\$) 1MHUA 2	
1712	9wga	A	26	114	1.1e-11	0.03	-0.20		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1715	1c0t	A	4	226	0	0.08	-0.20		HIV-1 REVERSE TRANSCRIPTASE (A-CHAIN); CHAIN: A; HIV-1 REVERSE TRANSCRIPTASE (B-CHAIN); CHAIN: B;	TRANSFERASE HIV-1 REVERSE TRANSCRIPTASE, AIDS, NON-NUCLEOSIDE INHIBITOR, 2 DRUG DESIGN
1715	1c9t	A	4	225	1.5e-99	0.22	-0.20		HIV-1 REVERSE TRANSCRIPTASE (CHAIN A); CHAIN: A; HIV-1 REVERSE TRANSCRIPTASE (CHAIN B); CHAIN: B; ANTIBODY (LIGHT	TRANSFERASE/IMMUNE SYSTEM/DNA HIV-1 RT, HIV-1 RT, HIV, REVERSE TRANSCRIPTASE, MET184ILE, 3TC, PROTEIN-DNA 2 COMPLEX, DRUG RESISTANCE, M184I, TRANSFERASE/IMMUNE 3 SYSTEM/DNA

1012

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL ID score	Compound	PDB annotation
									CHAIN: CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H; DNA (5'-CHAIN: T; DNA (5'-CHAIN: P;	
1715	1har		18	129	3e-61	0.02	-0.20		REVERSE TRANSCRIPTASE HIV-1 REVERSE TRANSCRIPTASE (AMINO-TERMINAL HALP) (FINGERS 1HAR 3 AND PALM SUBDOMAINS) (RT216) (E.C.2.7.7.49) 1HAR 4	
1715	1vrt	A	4	226	0	0.02	-0.20		HIV-1 REVERSE TRANSCRIPTASE; 1VRT 4 CHAIN: A, B; 1VRT 5	NUCLEOTIDYLTRANSFERASE HIV-1 RT; 1VRT 6 HIV-1 REVERSE TRANSCRIPTASE 1VRT 15
1715	1vrt	B	6	226	0	0.03	-0.20		HIV-1 REVERSE TRANSCRIPTASE; 1VRT 4 CHAIN: A, B; 1VRT 5	NUCLEOTIDYLTRANSFERASE HIV-1 RT; 1VRT 6 HIV-1 REVERSE TRANSCRIPTASE 1VRT 15
1718	1dix	A	222	253	0.0059	0.63	-0.20		SUBSTRATE BINDING DOMAIN OF DNK; CHAIN: A; SUBSTRATE PEPTIDE (7 RESIDUES); CHAIN: B;	COMPLEX (MOLECULAR CHAPERONE/PEPTIDE) DNK, HEAT SHOCK PROTEIN 70 KDA (HSP70), COMPLEX 2 (MOLECULAR CHAPERONE/PEPTIDE)
1718	1e3p	A	222	264	0.00041	0.24	-0.20		GUANOSINE PENTAPHOSPHATE SYNTHETASE; CHAIN: A;	POLYRIBONUCLEOTIDE TRANSFERASE POLYNUCLEOTIDE PHOSPHORYLASE, GUANOSINE POLYRIBONUCLEOTIDE TRANSFERASE, ATP-GTP DIPHOSPHOTRANSFERASE, 2 RNA PROCESSING, RNA DEGRADATION

1013

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1718	1edh	A	363	544	3e-25	0.26	-0.20		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1718	1edh	A	373	544	1.2e-26	-0.00	-0.20		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1718	1edh	A	52	160	1.3e-18	0.19	-0.20		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1718	1edh	A	65	278	3e-51			88.39	E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1718	1efu	B	214	243	0.00041	0.35	-0.20		ELONGATION FACTOR TU; CHAIN: A, C; ELONGATION FACTOR TS; CHAIN: B, D;	COMPLEX (TWO ELONGATION FACTORS) ELONGATION FACTOR FOR TRANSFER, HEAT UNSTABLE, ELONGATION FACTOR FOR TRANSFER, HEAT STABLE, ELONGATION FACTOR, COMPLEX (TWO ELONGATION FACTORS)
1718	1efu	B	220	267	0.0012	0.11	-0.20		ELONGATION FACTOR TU; CHAIN: A, C; ELONGATION FACTOR	COMPLEX (TWO ELONGATION FACTORS) ELONGATION FACTOR FOR TRANSFER, HEAT

1014

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									TS; CHAIN: B, D;	UNSTABLE, ELONGATION FACTOR FOR TRANSFER, HEAT STABLE, ELONGATION FACTOR, COMPLEX (TWO ELONGATION FACTORS)
1718	1g5z	A	217	243	0.0059	0.04	-0.20		OUTER SURFACE PROTEIN C; CHAIN: A;	IMMUNE SYSTEM SURFACE PROTEIN, ALPHA HELIX PROTEIN
1718	1g5z	A	218	243	0.0059	1.35	-0.20		OUTER SURFACE PROTEIN C; CHAIN: A;	IMMUNE SYSTEM SURFACE PROTEIN, ALPHA HELIX PROTEIN
1718	1ncg		63	158	3e-17	0.53	-0.20		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN
1718	1ncg		64	158	1.8e-28	0.45	-0.20		N-CADHERIN; INCG 3	CADHERIN INCG 13
1718	1ncg		6	101	7.5e-21			51.82	N-CADHERIN; INCG 3	CELL ADHESION PROTEIN
1718	1nci	B	63	160	1.3e-16	0.61	-0.20		N-CADHERIN; INCI 3	CELL ADHESION PROTEIN
1718	1nci	B	64	160	2.4e-27	0.38	-0.20		N-CADHERIN; INCI 3	CADHERIN INCI 13
1718	1ncj	A	370	544	3e-24	0.29	-0.20		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN
1718	1ncj	A	52	160	1.5e-16	0.21	-0.20		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN
1718	1ncj	A	64	285	1.5e-54			91.54	N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN
1718	1ncj	A	64	302	1.5e-54	0.05	-0.20		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN
1718	1suh		63	164	9e-20	0.63	-0.20		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1718	1suh		64	163	5.9e-24	0.42	-0.20		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1718	1suh		6	107	1e-22			54.43	EPITHELIAL CADHERIN;	CELL ADHESION UVOMORULIN;

1015

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1722	1bu7	A	110	531	4.2e-84	0.59	1.00		CHAIN: NULL;	CADHERIN, CALCIUM BINDING, CELL ADHESION
									CYTOCROME P450;	OXIDOREDUCTASE FATTY ACID HYDROXYLASE; FATTY ACID MONOOXYGENASE, HEMOPROTEIN, P450 REMARK
1722	1bu7	A	89	533	4.2e-84			185.73	CYTOCROME P450; CHAIN: A, B;	OXIDOREDUCTASE FATTY ACID HYDROXYLASE; FATTY ACID MONOOXYGENASE, HEMOPROTEIN, P450 REMARK
1722	1cpt		110	530	5.6e-28	0.19	1.00		OXIDOREDUCTASE(OX YGENASE) CYTOCHROME P450- TERP 1CPT 3	
1722	1cpt		275	530	9.6e-40	0.06	0.69		OXIDOREDUCTASE(OX YGENASE) CYTOCHROME P450- TERP 1CPT 3	
1722	1cpt		92	531	5.6e-28			103.94	OXIDOREDUCTASE(OX YGENASE) CYTOCHROME P450- TERP 1CPT 3	
1722	1dt6	A	92	502	2.8e-74	0.22	1.00		CYTOCROME P450 2C5; CHAIN: A;	OXIDOREDUCTASE PROGESTERONE 21-HYDROXYLASE, CYP11C5 P450 1, MEMBRANE PROTEIN, PROGESTERONE 21-HYDROXYLASE, BENZO(A) 2 PYRENE HYDROXYLASE, ESTRADIOL 2-HYDROXYLASE, P450, CYP2C5
1722	1dt6	A	92	533	2.8e-74			140.71	CYTOCROME P450 2C5; CHAIN: A;	OXIDOREDUCTASE PROGESTERONE 21-HYDROXYLASE, CYP11C5 P450 1,

1016

Table 5

SEQ ID No:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1722	1e9x	A	110	532	5.6e-76	0.61	1.00		CYTOCHROME P450 51-LIKE RV0764C; CHAIN: A;	MEMBRANE PROTEIN, PROGESTERONE 21-HYDROXYLASE, BENZO(A) 2 PYRENE HYDROXYLASE, ESTRADIOL 2-HYDROXYLASE, P450, CYP2C5
1722	1e9x	A	111	532	3.6e-94	0.72	1.00		CYTOCHROME P450 51-LIKE RV0764C; CHAIN: A;	OXIDOREDUCTASE CYP51, 14ALPHA-STEROL DEMETHYLASE; CYTOCHROME P450, 14 ALPHA-STEROL DEMETHYLASE, AZOLE 2 INHIBITORS
1722	1e9x	A	1	257	2.8e-45			79.09	CYTOCHROME P450 51-LIKE RV0764C; CHAIN: A;	OXIDOREDUCTASE CYP51, 14ALPHA-STEROL DEMETHYLASE; CYTOCHROME P450, 14 ALPHA-STEROL DEMETHYLASE, AZOLE 2 INHIBITORS
1722	1e9x	A	78	532	3.6e-94			308.79	CYTOCHROME P450 51-LIKE RV0764C; CHAIN: A;	OXIDOREDUCTASE CYP51, 14ALPHA-STEROL DEMETHYLASE; CYTOCHROME P450, 14 ALPHA-STEROL DEMETHYLASE, AZOLE 2 INHIBITORS
1722	1eup	A	122	501	1.2e-47	0.46	1.00		CYTOCHROME P450ERYF; CHAIN: A	OXIDOREDUCTASE CYTOCHROME P450, STEROID, ANDROSTENEDIONE, CYTOCHROME 2 P450ERYF
1722	1eup	A	124	500	7e-37	0.37	1.00		CYTOCHROME	OXIDOREDUCTASE



1017

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									P450ERYF; CHAIN: A	CYTOCHROME P450, STEROID, ANDROSTENEDIONE, CYTOCHROME 2 P450ERYF
1722	1eup	A	91	530	1.2e-47			108.69	CYTOCHROME P450ERYF; CHAIN: A	OXIDOREDUCTASE CYTOCHROME P450, STEROID, ANDROSTENEDIONE, CYTOCHROME 2 P450ERYF
1722	1f24	A	106	526	4.8e-42			102.76	NITRIC OXIDE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE NITRIC OXIDE REDUCTASE, CYTOCHROME P450NOR
1722	1f24	A	122	498	4.8e-42	0.21	0.72		NITRIC OXIDE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE NITRIC OXIDE REDUCTASE, CYTOCHROME P450NOR
1722	1f24	A	135	500	5.6e-21	0.16	1.00		NITRIC OXIDE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE NITRIC OXIDE REDUCTASE, CYTOCHROME P450NOR
1722	1fgx	A	8	318	0			342.49	BETA 1,4 GALACTOSYLTRANSFERASE; CHAIN: A, B;	TRANSFERASE BETA4GALTI; NUCLEOTIDE BINDING PROTEIN, ALPHA BETA ALPHA FOLD
1722	1f63	A	1	278	2.8e-52			54.01	ALPHA1,2-MANNOSIDASE; CHAIN: A;	HYDROLASE ALPHA-ALPHA7 BARREL
1722	1io7	A	107	530	1.1e-19			88.39	CYTOCHROME P450 CYP119; CHAIN: A, B;	OXIDOREDUCTASE THERMOPHILIC, CYTOCHROMO P450, CRYSTAL STRUCTURE
1722	1io7	A	277	502	7.2e-36	-0.30	0.04		CYTOCHROME P450 CYP119; CHAIN: A, B;	OXIDOREDUCTASE THERMOPHILIC, CYTOCHROMO P450, CRYSTAL STRUCTURE
1722	1qmq	A	121	525	2.4e-35	0.55	1.00		CYTOCHROME P450; CHAIN: A;	OXIDOREDUCTASE CAMPHOR 5-MONOXYGENASE
1722	1qmq	A	88	533	2.4e-35			90.48	CYTOCHROME P450; CHAIN: A;	OXIDOREDUCTASE CAMPHOR 5-MONOXYGENASE

1018

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1725	1cdy		35	117	0.00036	0.28	0.19		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: NULL;	OXIDOREDUCTASE(OXYGENASE), RU-SUBSTRATE, T-CELL SURFACE GLYCOPROTEIN IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC, LIPOPROTEIN, T-CELL SURFACE GLYCOPROTEIN
1725	1cvs	D	35	149	0.00048	-0.04	0.04		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1725	1eaj	A	35	130	2.4e-06	0.15	0.30		COXSACKIE VIRUS AND ADENOVIRUS RECEPTOR; CHAIN: A, B;	VIRUS/VIRAL PROTEIN RECEPTOR COXSACKIEVIRUS B-ADENOVIRUS RECEPTOR, HCAR, VIRUS/VIRAL PROTEIN RECEPTOR, IMMUNOGLOBULIN V DOMAIN FOLD, 2 SYMMETRIC DIMER
1725	1evt	C	34	134	0.00096	0.16	0.72		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1, FGFR1, IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1725	1iil	G	35	117	0.0096	0.14	0.16		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2, HBGF-2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE

1019

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	START AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1726	1eqk	A	11	105	1.1e-26			52.74	CH3 DOMAIN OF MAK33 ANTIBODY; CHAIN: A, B;	DOMAIN, B-TREFOIL
1726	1eqk	A	65	173	7e-29			52.57	CH3 DOMAIN OF MAK33 ANTIBODY; CHAIN: A, B;	IMMUNE SYSTEM CONSTANT DOMAIN, C1-SUBSET, IMMUNOGLOBULIN, IMMUNE SYSTEM
1726	1dn2	A	68	269	5.6e-51			53.86	IMMUNOGLOBULIN LAMBDA HEAVY CHAIN; CHAIN: A, B; ENGINEERED PEPTIDE; CHAIN: E, F;	IMMUNE SYSTEM FC IGG PHAGE DISPLAY PEPTIDE
1726	1e4k	A	60	269	2.8e-50			55.13	LOW AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C, FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
1726	1es0	B	1	167	1.4e-29			56.11	H-2 CLASS II HISTOCOMPATIBILITY ANTIGEN; CHAIN: A; 65 KD GLUTAMIC ACID DECARBOXYLASE+H-2 CLASS II CHAIN: B;	IMMUNE SYSTEM HISTOCOMPATIBILITY ANTIGEN, CLASS II MHC I-A(G7)
1726	1f3j	B	1	173	9.8e-33			50.10	H-2 CLASS II HISTOCOMPATIBILITY ANTIGEN; CHAIN: A, D; MHC CLASS II NOD; CHAIN: B, E; LYSOZYME C; CHAIN: P, Q;	IMMUNE SYSTEM HISTOCOMPATIBILITY ANTIGEN, MHC, PEPTIDE COMPLEX
1726	1fe2	D	1	170	2.8e-50			54.74	IMMUNOGLOBULIN IMMUNOGLOBULIN FC	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									AND FRAGMENT B OF PROTEIN A COMPLEX 1FC2 4	
1726	1fhe	A	1	169	2.8e-14			52.32	MHC CLASS II I-EK, ALPHA CHAIN; CHAIN: A, C; MHC CLASS II I-EK, BETA CHAIN; CHAIN: B, D;	IMMUNE SYSTEM HISTOCOMPATIBILITY ANTIGEN, MHC, PEPTIDE, IMMUNE SYSTEM
1726	1fv1	A	1	169	4.2e-17			53.07	MAJOR HISTOCOMPATIBILITY COMPLEX ALPHA CHAIN; CHAIN: A, D; MAJOR HISTOCOMPATIBILITY COMPLEX BETA CHAIN; CHAIN: B, E; MYELIN BASIC PROTEIN; CHAIN: C, F;	IMMUNE SYSTEM MHC CLASS II DR2A
1726	1hdm	B	7	175	5.6e-24			51.34	CLASS II HISTOCOMPATIBILITY ANTIGEN, M ALPHA CHAIN: A; CLASS II HISTOCOMPATIBILITY ANTIGEN, M BETA CHAIN: B;	IMMUNE SYSTEM RING6, HLA-DMA; RING7, HLA-DMB; HISTOCOMPATIBILITY PROTEIN, IMMUNE SYSTEM
1726	1ilc	A	1	169	2.8e-40			54.94	IG GAMMA-2A CHAIN C REGION; CHAIN: A, B;	IMMUNE SYSTEM IGG2A; IGG, FC
1726	1pfc		57	173	1.4e-27			54.04	IMMUNOGLOBULIN \$P/F\$C(PRIME) FRAGMENT OF AN IG*G1 1PFC 4	
1726	2iad	A	2	175	9.8e-14			51.43	MHC CLASS II I-AD; CHAIN: A, B;	MHC II MHC II, CLASS II MHC I-AD
1729	1de4	C	549	594	9.8e-05	-0.32	0.21		HEMOCHROMATOSIS	METAL TRANSPORT

1021

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									PROTEIN; CHAIN: A, D, G, BETA-2-MICROGLOBULIN; CHAIN: B, E, H; TRANSFERRIN RECEPTOR; CHAIN: C, F, I;	INHIBITOR/RECEPTOR HFE, HFE, HEREDITARY HEMOCHROMATOSIS, MHC CLASS I, TRANSFERRIN 2 RECEPTOR
1729	1dl2	A	157	260	2.8e-13	-0.45	0.19		CLASS I ALPHA-1,2-MANNOSIDASE; CHAIN: A;	HYDROLASE ALPHA-ALPHA HELIX BARREL
1729	1dl2	A	181	266	1.2e-19	-0.19	0.62		CLASS I ALPHA-1,2-MANNOSIDASE; CHAIN: A;	HYDROLASE ALPHA-ALPHA HELIX BARREL
1729	1dl2	A	194	374	5.6e-41	-0.21	0.10		CLASS I ALPHA-1,2-MANNOSIDASE; CHAIN: A;	HYDROLASE ALPHA-ALPHA HELIX BARREL
1729	1fo3	A	120	365	1.2e-54	0.06	1.00		ALPHA1,2-MANNOSIDASE; CHAIN: A;	HYDROLASE ALPHA-ALPHA7 BARREL
1729	1fo3	A	159	269	2.8e-19	-0.01	0.27		ALPHA1,2-MANNOSIDASE; CHAIN: A;	HYDROLASE ALPHA-ALPHA7 BARREL
1731	1f88	A	7	327	4.2e-96			122.22	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1731	1f88	A	8	345	4.2e-96	0.05	0.21		RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1731	1f88	B	2	314	5.6e-90			104.29	RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1731	1f88	B	8	312	5.6e-90	-0.13	0.03		RHODOPSIN; CHAIN: A, B	COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1732	1a4k	B	36	256	1.1e-75			84.51	ANTIBODY FAB; CHAIN: L, H, A, B;	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
1732	1ae6	H	35	260	1.4e-89	0.10	0.81		ANTIBODY CTM01; CHAIN: L, H;	IMMUNOGLOBULIN, ANTIBODY, CATALYTIC ANTIBODY, DIELS ALDER, 2 GERMLINE
1732	1afv	H	35	263	5.6e-88	-0.02	0.95		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	IMMUNOGLOBULIN, FAB FRAGMENT, HUMANSATISATION
1732	1a1l	H	36	244	1.1e-75			84.63	FAB59.1; CHAIN: L, H; AIB142; CHAIN: P;	COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN) HIV-1 CA, HIV CA, HIV P24, P24; FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN), HIV, CAPSID PROTEIN, 2 P24
1732	1a1f	A	154	287	9.8e-49	0.54	0.88		ANTL-IDIOTYPIC FAB 409.5.3 (IGG2A) FAB; CHAIN: A, B, L, H	COMPLEX (ANTIBODY/PEPTIDE) COMPLEX (ANTIBODY/PEPTIDE), ANTIBODY, CONSTRAINED HIV-1 V3 2 LOOP PEPTIDE, IMMUNOGLOBULIN
1732	1a1f	A	34	256	2.8e-34			86.97	ANTL-IDIOTYPIC FAB 409.5.3 (IGG2A) FAB; CHAIN: A, B, L, H	IMMUNOGLOBULIN IMMUNOGLOBULIN, C REGION, V REGION
1732	1b2w	L	154	287	1.4e-49	0.13	0.68		ANTIBODY (LIGHT	IMMUNOGLOBULIN, C REGION, V REGION

IMMUNE SYSTEM

1023

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRUCTURE, GAMMA-3 INTERFERON, IMMUNE SYSTEM
1732	1b6d	A	154	287	1.3e-48	0.34	0.76		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
1732	1bj1	J	154	287	2.8e-50	0.26	0.71		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
1732	1c5d	B	36	248	1.4e-68			85.03	MONOCLONAL ANTIBODY AGAINST THE MAIN IMMUNOGENIC CHAIN: L, A; MONOCLONAL ANTIBODY AGAINST THE MAIN IMMUNOGENIC CHAIN: H, B;	IMMUNE SYSTEM IMMUNOGLOBULIN
1732	1ce1	H	36	255	8.4e-74			89.67	CAMPATH-1H; LIGHT CHAIN; CHAIN: L; CAMPATH-1H; HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
1732	1cf8	H	36	247	2.8e-77			87.27	CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, TERPENOID SYNTHASE,

1024

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;	CARBOCATTON, 2 CYCLIZATION CASCADE
1732	1e18	L	154	287	9.8e-49	0.28	0.74		CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, TERPENOID SYNTHASE, CARBOCATTON, 2 CYCLIZATION CASCADE
1732	1e1c	B	35	227	1.4e-85	-0.06	0.77		IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
1732	1dee	A	154	287	2.8e-50	0.25	0.68		IGM RF 2A2; CHAIN: A; C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
1732	1dqd	H	36	246	5.6e-78			91.69	FAB HGR-2 F6; CHAIN: L; FAB HGR-2 F6; CHAIN: H;	IMMUNE SYSTEM GLUCAGON RECEPTOR, MONOCLONAL ANTIBODY, FAB, RECEPTOR 2 ANTAGONIST, TYPICAL IMMUNOGLOBULIN FOLD, LIGHT CHAIN, HEAVY 3 CHAIN, ANTIGEN BINDING SITE, COMPLEMENTARITY-DETERMINING 4 REGIONS
1732	1dqq	B	36	251	1.4e-72			85.70	ANTI-LYSOZYME ANTIBODY HYHEL-63 (LIGHT CHAIN); CHAIN: H;	IMMUNE SYSTEM ANTI-LYSOZYME ANTIBODY, HYHEL-63, HEN EGG WHITE LYSOZYME



1025

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A, C; ANTI-LYSOZYME ANTIBODY HYHEL-63 (HEAVY CHAIN); CHAIN: B, D;	
1732	1dzb	A	35	260	7e-96	0.44	0.99		SCFV FRAGMENT 1F9; CHAIN: A, B; TURKEY EGG-WHITE LYSOZYME C; CHAIN: X, Y;	COMPLEX (ANTIBODY ANTIGEN) 1,4-BETA-N-ACETYLMURAMIDASE C; SINGLE-DOMAIN ANTIBODY, TURKEY EGG-WHITE LYSOZYME, 2 ANTIBODY-PROTEIN COMPLEX, SINGLE-CHAIN FV FRAGMENT
1732	1dzb	A	36	259	7e-96			95.55	SCFV FRAGMENT 1F9; CHAIN: A, B; TURKEY EGG-WHITE LYSOZYME C; CHAIN: X, Y;	COMPLEX (ANTIBODY ANTIGEN) 1,4-BETA-N-ACETYLMURAMIDASE C; SINGLE-DOMAIN ANTIBODY, TURKEY EGG-WHITE LYSOZYME, 2 ANTIBODY-PROTEIN COMPLEX, SINGLE-CHAIN FV FRAGMENT
1732	1e6o	H	36	263	4.2e-88	0.14	0.95		IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L; IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H;	IMMUNOGLOBULIN FAB, ANTIBODY, ANTIGEN, HIV-1, P24, CA
1732	1eap	B	36	221	2.8e-80	-0.14	0.24		CATALYTIC ANTIBODY 17E8 COMPLEXED WITH PHENYL [1-(1-N-SUCCINYLAMINO)PENTYL] 1EAP 3 PHOSPHONATE 1EAP 4	
1732	1egj	H	36	227	7e-83	-0.06	0.47		CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR; CHAIN: A; ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY	IMMUNE SYSTEM CYTOKINE RECEPTOR COMPLEXED TO AN ANTIBODY

1026

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1732	1f3r	B	36	273	2.8e-84			99.76	CHAIN: H ACETYLCHOLINE RECEPTOR ALPHA; CHAIN: A; FV ANTIBODY FRAGMENT; CHAIN: B;	IMMUNE SYSTEM IG-FOLD, IMMUNO COMPLEX, ANTIBODY-ANTIGEN, BETA-TURN
1732	1f58	H	36	246	4.2e-79			86.22	IGG1 ANTIBODY 58.2 (LIGHT CHAIN); CHAIN: L; IGG1 ANTIBODY 58.2 (HEAVY CHAIN); CHAIN: H; EXTERIOR MEMBRANE GLYCOPROTEIN(GP120); CHAIN: P;	IMMUNE SYSTEM FAB 58.2; FAB 58.2; V3 LOOP; IMMUNOGLOBULIN, FAB, HIV-1, GP120, V3, IMMUNE SYSTEM
1732	1f8t	H	36	246	2.8e-79			86.44	ANTIBODY FAB FRAGMENT (LIGHT CHAIN); CHAIN: L; ANTIBODY FAB FRAGMENT (HEAVY CHAIN); CHAIN: H	IMMUNE SYSTEM MONOCLONAL ANTIBODY, ANTIGEN-BINDING FRAGMENT, INTERLEUKIN-2, X-RAY ANALYSIS, CRYSTAL
1732	1fai	H	35	229	5.6e-84	-0.05	0.49		IMMUNOGLOBULIN FAB FRAGMENT FROM A MONOCLONAL ANTILARSONATE ANTIBODY, R19.9 IFAI 3 (IGG2B.KAPPA) IFAI 4	
1732	1fbi	H	35	263	2.8e-88	0.08	0.72		COMPLEX (ANTIBODY/ANTIGEN) FAB FRAGMENT OF THE MONOCLONAL ANTIBODY F9.13.7 (IGG1) IFAI 3 COMPLEXED WITH LYSOZYME (E.C.3.2.1.17) IFAI 4	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1732	1f1g	L	34	256	2.8e-35			85.48	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (KAPPA LIGHT CHAIN) FAB FRAGMENT 1FIG 3	
1732	1fvd	A	154	287	2.8e-49	0.20	0.60		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
1732	1fvd	B	34	264	1.4e-80			89.40	IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
1732	1gpo	H	36	247	2.8e-73			86.78	ANTIBODY M41; CHAIN: L, H, M, I;	IMMUNOGLOBULIN PROTEIN ENGINEERING, ANTIBODY DESIGN, IMMUNOGLOBULIN 2 STRUCTURE, ANTIGEN-BINDING SITE, CANONICAL CONFORMATION, 3 COMPLEMENTARITY- DETERMINING REGION
1732	1iai	H	35	227	9.8e-83	0.10	0.74		IDIOtypic FAB 730.1.4 (IGG1) OF VIRUS 1IAI 5 CHAIN: L, H; 1IAI 7 ANTI-IDIOtypic FAB 409.5.3 (IGG2A); 1IAI 9 CHAIN: M, I 1IAI 10	COMPLEX (IMMUNOGLOBULIN IGG1/IGG2A)
1732	1iai	H	36	253	9.8e-83			86.17	IDIOtypic FAB 730.1.4 (IGG1) OF VIRUS 1IAI 5 CHAIN: L, H; 1IAI 7 ANTI-IDIOtypic FAB 409.5.3 (IGG2A); 1IAI 9 CHAIN: M, I 1IAI 10	COMPLEX (IMMUNOGLOBULIN IGG1/IGG2A)
1732	1i1g	B	37	245	7e-81			84.57	IMMUNOGLOBULIN FAB (IGG2A, KAPPA) FRAGMENT (26-10)	

1028

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									COMPLEX WITH DIGOXIN 1IGJA 1 1IGJA 2	
1732	1igt	B	36	292	8.4e-93	0.05	0.36		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1732	1lmk	A	36	260	9.8e-91	0.49	1.00		IMMUNOGLOBULIN ANTI- PHOSPHATIDYLINOSITOL SPECIFIC PHOSPHOLIPASE C DIABODY 1LMK 3 SYNONYMS: 15MK16 DIABODY, SINGLE-CHAIN FV DIMER 1LMK 4	
1732	1lmk	A	37	260	9.8e-91			101.55	IMMUNOGLOBULIN ANTI- PHOSPHATIDYLINOSITOL SPECIFIC PHOSPHOLIPASE C DIABODY 1LMK 3 SYNONYMS: 15MK16 DIABODY, SINGLE-CHAIN FV DIMER 1LMK 4	
1732	1ma m	H	36	229	9.8e-79	0.12	0.60		IMMUNOGLOBULIN ANTIGEN-BINDING FRAGMENT (FAB) (GG2B, KAPPA) 1MAM 3	
1732	1mfa		36	260	7e-50			84.82	IMMUNOGLOBULIN FV FRAGMENT (MURINE SE155-4) COMPLEX WITH THE TRISACCHARIDE: 1MEA	

1029

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									3 ALPHA-D-GALACTOSE(1-2)/ALPHA-D-ABEQUOSE(1-3)/ALPHA-1MFA 4 D-MANNOSE (P1-OME) (PART OF THE CELL-SURFACE CARBOHYDRATE 1MFA 5 OF PATHOGENIC SALMONELLA) 1MFA 6	
1732	1mrd	H	36	229	1.4e-83	-0.02	0.78		IMMUNOGLOBULIN IGG JEL 103 FAB FRAGMENT COMPLEXED WITH 1MRD 3 INOSINE-5'-DIPHOSPHATE 1MRD 4	
1732	1ngp	H	35	263	2.8e-89	0.06	0.60		NI G9 (IGG1=LAMBDA=); CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN,
1732	1nqb	A	36	261	2.8e-98	0.51	0.99		SINGLE-CHAIN ANTIBODY FRAGMENT; CHAIN: A, C;	HEAVY (VH) DOMAIN, VARIABLE LIGHT (VL) ANTIBODY FRAGMENT, MULTIVALENT ANTIBODY, DOMAIN 2 SWAPPING, IMMUNOGLOBULIN
1732	1nqb	A	37	261	2.8e-98			95.40	SINGLE-CHAIN ANTIBODY FRAGMENT; CHAIN: A, C;	IMMUNOGLOBULIN VARIABLE HEAVY (VH) DOMAIN, VARIABLE LIGHT (VL) ANTIBODY FRAGMENT, MULTIVALENT ANTIBODY, DOMAIN 2 SWAPPING, IMMUNOGLOBULIN
1732	1osp	H	36	242	8.4e-72			96.00	FAB 184.1; CHAIN: L, H; OUTER SURFACE PROTEIN A; CHAIN: O;	COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN) OSP A; COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN), OUTER SURFACE 2 PROTEIN A COMPLEXED WITH FAB184.1,

1030

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										BORRELLIA BURGDORFERI 3 STRAIN B31
1732	1plg	H	35	227	7e-87	-0.18	0.71		IGG2A=KAPPA=, 1PLG 4 CHAIN: L, H, 1PLG 5	IMMUNOGLOBULIN
1732	1qkz	H	36	227	2.8e-82	0.11	0.42		ANTIBODY; CHAIN: H, L, PROTEIN G-PRIME; CHAIN: A; MAJOR OUTER MEMBRANE PROTEIN P1.16; CHAIN: P,	IMMUNE SYSTEM FAB, PORA, NEISSERIA MENINGITIDIS, PORIN
1732	1qok	A	35	260	2.8e-94	0.40	0.89		MFE-23 RECOMBINANT ANTIBODY FRAGMENT; CHAIN: A;	IMMUNOGLOBULIN
1732	1qok	A	36	260	2.8e-94			95.49	MFE-23 RECOMBINANT ANTIBODY FRAGMENT; CHAIN: A;	IMMUNOGLOBULIN, SINGLE-CHAIN FV, ANTI-CARCINOEMBRYONIC 2 ANTIGEN
1732	1sbs	L	154	287	7e-48	0.20	0.47		MONOCLONAL ANTIBODY 3A2; CHAIN: H, L;	IMMUNOGLOBULIN, SINGLE-CHAIN FV, ANTI-CARCINOEMBRYONIC 2 ANTIGEN
1732	25c8	L	154	287	5.6e-50	0.42	0.69		IGG 5C8; CHAIN: L, H;	MONOCLONAL ANTIBODY MONOCLONAL ANTIBODY, FAB-FRAGMENT, REPRODUCTION
1732	2cgr	H	36	229	5.6e-81	-0.06	0.51		IMMUNOGLOBULIN IGG2B (KAPPA) FAB FRAGMENT COMPLEXED WITH ANTIGEN 2CGR 3 N-(P-CYANOPHENYL)-N'-(DIPHENYLEMETHYL) GUANIDINEACETIC ACID 2CGR 4	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
1732	2fgrw	L	154	287	5.6e-50	0.26	0.63		IMMUNOGLOBULIN FAB	

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 1H52 (HUH52-OZ FAB) 2FGW 4	
1732	32c2	B	36	245	4.2e-77			85.02	IGG1 ANTIBODY 32C2; CHAIN: A; IGG1 ANTIBODY 32C2; CHAIN: B;	IMMUNE SYSTEM FAB, ANTIBODY, AROMATASE, P450
1732	3hfm	H	36	246	8.4e-74			92.13	COMPLEX(ANTIBODY-ANTIGEN) IGG1 FAB FRAGMENT (HY/HEL-10) AND LYSOZYME (E.C.3.2.1.17) 3HFM 4 COMPLEX 3HFM 5	
1734	1dk6	A	51	492	7e-74			67.96	CYTOCROME P450 2C5; CHAIN: A;	OXIDOREDUCTASE PROGESTERONE 21-HYDROXYLASE, CYP11C5 P450 1, MEMBRANE PROTEIN, PROGESTERONE 21-HYDROXYLASE, BENZO(A) 2 PYRENE HYDROXYLASE, ESTRADIOL 2-HYDROXYLASE, P450, CYP2C5
1739	1dg3	A	1	472	7e-91			73.28	INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1; CHAIN: A;	SIGNALING PROTEIN GUANINE NUCLEOTIDE-BINDING PROTEIN 1; GBP, GTP HYDROLYSIS, GDP, GMP, INTERFERON INDUCED, DYNAMIN 2 RELATED, LARGE GTPASE FAMILY, SIGNALING PROTEIN
1739	1dg3	A	73	373	7e-91	-0.16	0.21		INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1; CHAIN: A;	SIGNALING PROTEIN GUANINE NUCLEOTIDE-BINDING PROTEIN 1; GBP, GTP HYDROLYSIS, GDP, GMP, INTERFERON INDUCED,

1032

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1739	1f5n	A	1	472	0			88.75	INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1; CHAIN: A;	DYNAMIN 2 RELATED, LARGE GTPASE FAMILY, SIGNALING PROTEIN
										SIGNALING PROTEIN GBP, GTP HYDROLYSIS, GDP, GMP, INTERFERON INDUCED, DYNAMIN 2 RELATED, LARGE GTPASE FAMILY. GMPPNP, GPPNHP.
1739	1f5n	A	65	373	0	0.15	0.74		INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1; CHAIN: A;	SIGNALING PROTEIN GBP, GTP HYDROLYSIS, GDP, GMP, INTERFERON INDUCED, DYNAMIN 2 RELATED, LARGE GTPASE FAMILY. GMPPNP, GPPNHP.
1742	1byw	A	371	402	0.0024	-0.24	0.05		HUMAN ERG POTASSIUM CHANNEL; CHAIN: A;	MEMBRANE PROTEIN PAS DOMAIN, POTASSIUM CHANNEL DOMAIN, MEMBRANE PROTEIN
1742	1d06	A	363	463	1.1e-11	0.55	0.74		NITROGEN FIXATION REGULATORY PROTEIN FIXL; CHAIN: A;	SIGNALING PROTEIN OXYGEN SENSOR, HISTIDINE KINASE, PAS, HIGH-RESOLUTION, TWO-2 COMPONENT SYSTEM
1742	1d06	A	469	540	4.8e-05	0.24	0.78		NITROGEN FIXATION REGULATORY PROTEIN FIXL; CHAIN: A;	SIGNALING PROTEIN OXYGEN SENSOR, HISTIDINE KINASE, PAS, HIGH-RESOLUTION, TWO-2 COMPONENT SYSTEM
1742	1d06	A	573	618	0.0043	0.22	0.69		NITROGEN FIXATION REGULATORY PROTEIN FIXL; CHAIN: A;	SIGNALING PROTEIN OXYGEN SENSOR, HISTIDINE KINASE, PAS, HIGH-RESOLUTION, TWO-2 COMPONENT SYSTEM
1742	1d06	A	574	616	8.4e-06	0.28	0.99		NITROGEN FIXATION REGULATORY PROTEIN FIXL; CHAIN: A;	SIGNALING PROTEIN OXYGEN SENSOR, HISTIDINE KINASE, PAS, HIGH-RESOLUTION, TWO-2 COMPONENT SYSTEM



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1742	1dp6	A	359	465	8.4e-11	0.65	0.75		FIXL PROTEIN; CHAIN: A;	OXYGEN STORAGE/TRANSPORT FIXL, HEME, PAS DOMAIN, SIGNAL TRANSDUCTION, HISTIDINE 2 KINASE
1742	1dp6	A	360	462	1.2e-09	0.40	0.80		FIXL PROTEIN; CHAIN: A;	OXYGEN STORAGE/TRANSPORT FIXL, HEME, PAS DOMAIN, SIGNAL TRANSDUCTION, HISTIDINE 2 KINASE
1742	1dp6	A	469	542	7.2e-05	0.48	0.42		FIXL PROTEIN; CHAIN: A;	OXYGEN STORAGE/TRANSPORT FIXL, HEME, PAS DOMAIN, SIGNAL TRANSDUCTION, HISTIDINE 2 KINASE
1742	1dp6	A	574	616	0.00012	0.82	0.34		FIXL PROTEIN; CHAIN: A;	OXYGEN STORAGE/TRANSPORT FIXL, HEME, PAS DOMAIN, SIGNAL TRANSDUCTION, HISTIDINE 2 KINASE
1742	1g28	A	579	614	0.0006	0.34	0.46		PHY3 PROTEIN; CHAIN: A, B, C, D;	SIGNALING PROTEIN, ELECTRON TRANSPORT PHOTOTROPIN, LOV, PAS FOLD, PHOTORECEPTOR, FLAVOPROTEIN, 2 FMN-BINDING DOMAIN, ALPHA-BETA STRUCTURE
1745	1cfe		30	176	4.2e-14	0.78	1.00		PATHOGENESIS-RELATED PROTEIN P14A; CHAIN: NULL;	PATHOGENESIS-RELATED PROTEIN PATHOGENESIS-RELATED LEAF PROTEIN 6, ETHYLENE PATHOGENESIS-RELATED PROTEIN, PR-1 PROTEINS, 2 PLANT DEFENSE
1745	1qnx	A	1	170	4.2e-22	0.29	0.87		VES V 5; CHAIN: A;	ALLERGEN ANTIGEN 5; ANTIGEN 5, ALLERGEN, VESPID VENOM
1745	1qnx	A	28	185	1.2e-37	0.33	1.00		VES V 5; CHAIN: A;	ALLERGEN ANTIGEN 5; ANTIGEN 5, ALLERGEN, VESPID VENOM
1749	1cw3	A	27	75	1.4e-18	-0.58	0.96		ALLERGEN BQU C 1; CHAIN: A;	ALLERGEN LIPOCALIN, BETA BARREL

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1749	1i06	A	42	75	2.4e-07	-0.37	0.94		MAJOR URINARY PROTEIN I; CHAIN: A;	TRANSPORT PROTEIN ALPHA-2U-GLOBULIN I; LIPOCALIN, BETA-BARREL, PHEROMONE
1749	2a2u	A	27	75	9.8e-16	-0.24	0.84		ALPHA-2U-GLOBULIN; CHAIN: A, B, C, D	LIPID BINDING PROTEIN A2U-GLOBULIN, LIPID BINDING PROTEIN
1753	1bii	A	148	511	5.6e-41	0.17	0.46		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1753	1bii	A	149	511	2.8e-51			101.99	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1753	1bii	A	247	611	2.8e-40	0.30	0.42		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1753	1bii	A	59	416	2.8e-51	0.13	-0.07		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
1753	1cfo		511	685	1.3e-12	0.39	-0.08		NEURAL ADHESION MOLECULE DROSOPHILA NEUROGLAN (CHYMOTRYPTIC FRAGMENT CONTAINING THE ICFB 3 TWO AMINO PROXIMAL FIBRONECTIN TYPE III REPEATS ICFB 4 (RESIDUES 610 - 814)) ICFB 5	
1753	1cs6	A	142	513	1.4e-59			126.96	AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1753	1cs6	A	148	512	1.4e-54	0.20	0.11		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										ADHESION
1753	1cs6	A	247	612	4.2e-50	0.19	0.93		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1753	1cs6	A	57	416	1.4e-59	0.12	0.17		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1753	1cvs	C	159	325	2.8e-29	0.10	0.00		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1753	1cvs	C	240	415	1.1e-34	0.32	-0.05		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1753	1cvs	D	159	325	2.8e-30	0.01	0.06		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1753	1cvs	D	240	415	5.6e-36	0.15	0.00		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1753	1cvs	D	325	512	8.4e-37	0.01	-0.12		FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH

1036

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1753	1evs	D	417	611	2.8e-32	0.03	-0.11		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1753	1evs	D	58	235	4.2e-32	0.15	-0.11		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1753	1dgi	R	243	511	2.4e-21	-0.09	0.05		POLIOVIRUS RECEPTOR; CHAIN: R; VP1; CHAIN: 1; VP2; CHAIN: 2; VP3; CHAIN: 3; VP4; CHAIN: 4;	VIRUS/VIRAL PROTEIN, RECEPTOR CD155, PVR, HUMAN POLIOVIRUS, ELECTRON MICROSCOPY, 2 POLIOVIRUS-RECEPTOR COMPLEX, VIRUS/VIRAL PROTEIN, RECEPTOR
1753	1dh2	A	330	500	1.4e-25	-0.07	0.22		IMMUNOGLOBULIN LAMBDA HEAVY CHAIN; CHAIN: A, B; ENGINEERED PEPTIDE; CHAIN: E, F;	IMMUNE SYSTEM FC IGG PHAGE DISPLAY PEPTIDE
1753	1dx5	I	2	121	4.2e-17			55.65	THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II;

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
1753	1e4k	A	330	500	1.1e-25	-0.04	0.01		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC RECEPTOR CHAIN: C; FC FRAGMENT OF HUMAN IGG1; CHAIN: A, B;	COMPLEX CD16; IGG1-FC COMPLEX, FC FRAGMENT, IGG, FC, RECEPTOR, CD16, GAMMA
1753	1eaj	A	33	145	0.0036	-0.03	0.03		COXSACKIE VIRUS AND ADENOVIRUS RECEPTOR; CHAIN: A, B;	VIRUS/VIRAL PROTEIN RECEPTOR COXSACKIEVIRUS B-ADENOVIRUS RECEPTOR, HCAR, VIRUS/VIRAL PROTEIN RECEPTOR, IMMUNOGLOBULIN V DOMAIN FOLD, 2 SYMMETRIC DIMER
1753	1emm		36	121	4.2e-20			60.14	FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1753	1epf	A	253	399	1.4e-21	0.29	0.36		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1753	1epf	A	335	495	2.8e-21	0.02	-0.12		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1753	1epf	A	424	595	1.1e-21	0.12	-0.02		NEURAL CELL ADHESION MOLECULE;	CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD,

1038

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1753	1epf	A	54	235	5.6e-26	0.23	-0.01		CHAIN: A, B, C, D; NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	GLYCOPROTEIN CELL ADHESION NCAM; NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1753	1ev2	E	243	415	1.3e-31	0.11	-0.18		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1753	1ev2	E	429	611	9.8e-31	0.27	-0.03		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1753	1ev2	G	243	419	2.8e-35	0.16	-0.08		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1753	1ev2	G	429	615	1.3e-33	0.02	-0.09		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
1753	1evt	C	159	325	5.6e-30	-0.05	0.06		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD

1039

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1753	1evt	C	240	415	4.2e-35	0.16	-0.14		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE 1-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1753	1evt	C	325	512	1.4e-35	0.05	-0.18		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE 1-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1753	1evt	C	417	611	1.4e-31	0.02	-0.11		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE 1-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1753	1f2q	A	247	420	2.8e-18	0.45	0.63		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RL-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1753	1f2q	A	330	515	7e-18	0.15	-0.11		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RL-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN
1753	1f6a	A	149	331	4.8e-23	-0.11	0.13		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1753	1f6a	A	247	419	4.2e-17	0.48	0.70		HIGH AFFINITY	IMMUNE SYSTEM HIGH AFFINITY

1040

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1753	1f97	A	155	320	1.4e-25	0.47	0.87		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1753	1f97	A	254	410	8.4e-29	0.26	-0.05		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1753	1f97	A	421	606	1.4e-26	0.24	0.05		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1753	1f97	A	59	231	4.2e-27	0.09	0.60		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1753	1fc2	D	330	500	1.1e-25	0.06	0.59		IMMUNOGLOBULIN AND FRAGMENT B OF PROTEIN A COMPLEX IFC2 4	
1753	1fcg	A	149	327	1.2e-22	0.26	0.77		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1753	1fcg	A	239	401	8.4e-17	0.16	0.88		FC RECEPTOR FC(GAMMA)RIIA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1753	1fcg	A	329	512	2.8e-19	0.52	-0.17		FC RECEPTOR FC(GAMMA)RLA; CHAIN: A;	IMMUNE SYSTEM, MEMBRANE PROTEIN CD32; FC RECEPTOR, IMMUNOGLOBULIN, LEUKOCYTE, CD32
1753	1ftl	A	239	401	2.4e-16	0.34	1.00		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1753	1hj7	A	39	121	1.4e-19			51.87	LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
1753	1hzh	H	259	614	5.6e-17	0.01	-0.09		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
1753	1hzh	H	62	404	1.4e-22	0.24	0.60		IMMUNOGLOBULIN HEAVY CHAIN; CHAIN: H, K; IMMUNOGLOBULIN LIGHT CHAIN; CHAIN: L, M;	IMMUNE SYSTEM IGG; IMMUNOGLOBULIN, ANTIBODY, B12
1753	1iic	A	332	506	1.4e-20	0.25	0.09		IG GAMMA-2A CHAIN C REGION; CHAIN: A, B;	IMMUNE SYSTEM IGG2A; IGG, FC
1753	1iir	A	348	611	8.4e-15	0.05	-0.08		INTERLEUKIN-6 RECEPTOR BETA CHAIN; CHAIN: A; VIRAL IL-6; CHAIN: B;	CYTOKINE GPI30; FUNCTIONAL INTERLEUKIN-6 HOMOLOG; CYTOKINE/RECEPTOR COMPLEX, GPI30, VIRAL IL-6, CRYSTAL 2 STRUCTURE
1753	1ie5	A	323	415	1.4e-15	0.10	-0.11		NEURAL CELL ADHESION MOLECULE;	CELL ADHESION N-CAM; INTERMEDIATE

1042

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: A;	IMMUNOGLOBULIN FOLD
1753	1igt	B	159	499	1.1e-25	-0.12	0.24		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1753	1igt	B	66	410	1.3e-22	0.15	-0.12		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
1753	1igy	B	64	410	5.6e-23	-0.22	0.19		IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
1753	1iil	G	159	329	5.6e-27	0.24	0.13		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF- 2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
1753	1iil	G	243	419	1.4e-34	0.26	0.04		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF- 2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
1753	1iil	G	429	615	9.8e-36	0.23	-0.08		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2, HBGF- 2, BASIC FIBROBLAST GROWTH FACTOR, FGFR2, KERATINOCYTE GROWTH FACTOR RECEPTOR; IMMUNOGLOBULIN LIKE DOMAIN, B-TREFOIL
1753	1iib	B	169	413	1.3e-16	0.10	-0.11		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD,

1043

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									RECEPTOR; CHAIN: B;	TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1753	1itb	B	352	609	2.8e-19	0.24	-0.15		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1753	1itb	B	39	326	3.6e-21	-0.03	0.12		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1753	1mco	H	123	511	4.2e-37			100.32	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION 1MCO 3	
1753	1mco	H	54	404	1.4e-26	-0.07	0.74		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION 1MCO 3	
1753	1nct		248	326	1.3e-16	0.05	0.13		TTTN; CHAIN: NULL; 3	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN

1044

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1753	1nct		519	612	2.8e-10	0.07	-0.19		TTTN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
1753	1neu		43	130	0.0072	-0.02	0.11		MYELIN P0 PROTEIN; CHAIN: NULL;	STRUCTURAL PROTEIN MYELIN, STRUCTURAL PROTEIN, GLYCOPROTEIN, TRANSMEMBRANE, PHOSPHORYLATION, IMMUNOGLOBULIN FOLD, SIGNAL, MYELIN 2 MEMBRANE ADHESION MOLECULE
1753	1qg3	A	487	615	1.2e-09	0.21	-0.15		INTEGRIN BETA-4 SUBUNIT; CHAIN: A, B;	STRUCTURAL PROTEIN INTEGRIN, HEMIDESMOSOME, FIBRONECTIN, CARCINOMA, STRUCTURAL 2 PROTEIN
1753	1qr4	A	487	610	1.2e-10	0.37	-0.08		TENASCIN; CHAIN: A, B;	STRUCTURAL PROTEIN TENASCIN, FIBRONECTIN TYPE-III, HEPARIN, EXTRACELLULAR 2 MATRIX, ADHESION, FUSION PROTEIN, STRUCTURAL PROTEIN
1753	1tmm		248	326	1.3e-16	0.40	0.11		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNM 4 1TNM 58	
1753	1tmm		519	612	2.8e-10	0.23	-0.17		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM3 (NMR, MINIMIZED AVERAGE STRUCTURE)	

1045

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1753	1wio	A	161	508	2.4e-22	0.04	-0.15		1TNM 4 1TNM 58 T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
1753	1wio	A	44	402	4.8e-24	0.19	-0.17		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
1753	2di	A	149	314	4.8e-21	-0.20	0.39		MHC CLASS I NK CELL RECEPTOR PRECURSOR; CHAIN: A;	IMMUNE SYSTEM P58 NATURAL KILLER CELL RECEPTOR, KIR, NATURAL KILLER RECEPTOR, INHIBITORY RECEPTOR, 2 IMMUNOGLOBULIN
1753	2fcb	A	239	401	9.6e-16	0.21	0.76		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1753	2fcb	A	329	514	5.6e-19	0.21	-0.15		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1753	2ncn		54	143	5.6e-09	0.09	-0.12		NEURAL CELL ADHESION MOLECULE; CHAIN: NULL;	CELL ADHESION NCAM DOMAIN 1; CELL ADHESION, GLYCOPROTEIN, HEPARIN- BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, SIGNAL
1753	3ncn	A	55	141	2.8e-08	-0.00	-0.19		NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM; CHAIN: A;	CELL ADHESION PROTEIN NCAM MODULE 2; CELL ADHESION, GLYCOPROTEIN, HEPARIN- BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE,

1046

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verity score	PMF score	SEQFOL D score	Compound	PDB annotation
1753	8fab	A	334	507	2.8e-16	-0.00	-0.09		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	IMMUNOGLOBULIN FOLD, HOMOPHILIC 3 BINDING, CELL ADHESION PROTEIN
1753	8fab	A	421	599	1.1e-15	0.06	-0.18		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
1757	1hd7	A	367	630	1.2e-14	-0.08	0.06		DNA-(APURINIC OR APYRIMIDINIC SITE) LYASE; CHAIN: A;	DNA REPAIR AP ENDONUCLEASE 1, HAP1, REF1, APE1; DNA REPAIR, ENDONUCLEASE, APE1, HAP1, REF-1
1757	1i9z	A	1	279	5.6e-86			109.20	PHOSPHATIDYLINOSITO L PHOSPHATE PHOSPHATASE; CHAIN: A;	HYDROLASE SYNAPTOJANIN; SPSYNAPTOJANIN, IPP5C, IP3, IP2,
1757	1i9z	A	345	702	1.3e-86			143.57	PHOSPHATIDYLINOSITO L PHOSPHATE PHOSPHATASE; CHAIN: A;	HYDROLASE SYNAPTOJANIN; SPSYNAPTOJANIN, IPP5C, IP3, IP2,
1757	1i9z	A	359	627	1.3e-86	0.62	1.00		PHOSPHATIDYLINOSITO L PHOSPHATE PHOSPHATASE; CHAIN: A;	HYDROLASE SYNAPTOJANIN; SPSYNAPTOJANIN, IPP5C, IP3, IP2,
1757	1i9z	A	363	621	4.8e-71	0.54	1.00		PHOSPHATIDYLINOSITO L PHOSPHATE PHOSPHATASE; CHAIN: A;	HYDROLASE SYNAPTOJANIN; SPSYNAPTOJANIN, IPP5C, IP3, IP2,

1047

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1757	1i9z	A	750	789	0.00098	-0.33	0.39		A; PHOSPHATIDYLINOSITOL PHOSPHATE PHOSPHATASE; CHAIN: A;	HYDROLASE SYNAPTOANIN; SPSYNAPTOANIN, IPP5C, IP3, IP2,
1757	1i9z	A	753	791	1.2e-09	-0.36	0.47		PHOSPHATIDYLINOSITOL PHOSPHATE PHOSPHATASE; CHAIN: A;	HYDROLASE SYNAPTOANIN; SPSYNAPTOANIN, IPP5C, IP3, IP2,
1758	1qdv	A	107	214	1.4e-28	0.18	0.12		KVL2 VOLTAGE-GATED POTASSIUM CHANNEL; CHAIN: A, B, C, D;	SIGNALING PROTEIN VOLTAGE-GATED POTASSIUM CHANNEL, TETRAMERIZATION DOMAIN, 2 INTRACELLULAR GATE, TETRAMER
1759	1esj	A	260	524	1.1e-16	0.45	0.23		HYDROXYETHYLTHIAZOLE KINASE; CHAIN: A, B, C;	TRANSFERASE THZ KINASE; TRIMER, ALPHA-BETA PROTEIN
1763	1a4y	A	60	364	2.4e-38	0.32	0.93		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1763	1a9n	A	152	306	1.2e-15	0.09	0.03		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C, U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	A	176	341	1.2e-14	0.42	-0.03		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C, U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	A	210	361	4.8e-13	0.49	0.23		U2 RNA HAIRPIN IV;	COMPLEX (NUCLEAR

1048

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	A	255	361	1.4e-06	0.20	0.60		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	A	63	169	5.6e-06	0.38	0.30		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	A	67	216	2.4e-22	0.47	0.33		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	A	91	234	3.6e-21	0.75	0.96		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	C	210	374	6e-13	0.26	-0.07		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	C	255	361	1.4e-06	-0.04	0.72		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	C	67	237	1.2e-23	0.44	0.51		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1763	1a9n	C	412	518	2.4e-16	0.27	0.11		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
1763	1cvs	C	417	519	2.4e-14	0.46	0.17		FIBROBLAST GROWTH	GROWTH FACTOR/GROWTH



1049

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1763	1evs	D	417	518	7.2e-15	0.38	0.10		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1763	1evs	D	420	502	6e-16	0.52	0.65		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
1763	1d0b	A	128	331	1.1e-21	0.13	0.37		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1763	1d0b	A	248	402	1.4e-20	0.16	0.98		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1763	1d0b	A	42	195	4.2e-19	0.45	0.90		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1763	1d0b	A	59	283	1.4e-18	-0.15	0.21		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1763	1d0b	A	61	236	1.1e-19	0.68	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION

1050

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1763	1d0b	A	63	211	2.4e-20	0.35	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1763	1d0b	A	80	309	2.4e-20	0.46	0.90		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1763	1dce	A	135	288	8.4e-06	-0.11	0.06		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1763	1dce	A	231	353	5.6e-13	0.24	0.78		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1763	1dce	A	255	360	7e-12	0.08	1.00		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1763	1dce	A	42	168	1.4e-07	0.09	-0.03		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-

1051

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1763	1ds9	A	216	363	1.1e-12	-0.01	0.10		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1763	1ds9	A	240	363	5.6e-15	-0.28	0.12		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1763	1epf	A	420	518	7.2e-14	0.51	0.39		NEURAL CELL ADHESION MOLECULE; CHAIN: A, B, C, D;	CELL ADHESION NCAM, NCAM, IMMUNOGLOBULIN FOLD, GLYCOPROTEIN
1763	1ev2	E	417	518	4.8e-15	0.38	0.39		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1763	1ev2	G	420	507	1.2e-15	0.17	0.15		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREEFOLD FOLD
1763	1l2q	A	411	518	2.4e-13	0.36	0.31		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A;	IMMUNE SYSTEM FC-EPSILON RI-ALPHA; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN

1052

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1763	1f6a	A	411	518	2.4e-14	0.23	0.22		HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR CHAIN: A; IG EPSILON CHAIN C REGION; CHAIN: B, D;	IMMUNE SYSTEM HIGH AFFINITY IGE-FC RECEPTOR, FC(EPSILON) IGE-FC; IMMUNOGLOBULIN FOLD, GLYCOPROTEIN, RECEPTOR, IGE-BINDING 2 PROTEIN, IGE ANTIBODY, IGE-FC
1763	1f97	A	411	518	8.4e-14	0.27	0.35		JUNCTION ADHESION MOLECULE; CHAIN: A;	CELL ADHESION IMMUNOGLOBULIN SUPERFAMILY, BETA-SANDWICH FOLD
1763	1f1l	A	411	518	4.8e-14	0.42	0.01		LOW AFFINITY IMMUNOGLOBULIN GAMMA FC REGION CHAIN: A;	IMMUNE SYSTEM RECEPTOR BETA SANDWICH, IMMUNOGLOBULIN-LIKE, RECEPTOR
1763	1f6l	A	250	313	5.6e-07	-0.48	0.24		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1763	1f6l	A	78	138	7.2e-07	-0.27	0.33		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1763	1f6l	B	232	288	1.4e-05	0.01	0.03		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1763	1f6l	B	250	313	5.6e-07	-0.27	0.22		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1763	1f6l	B	65	113	3.6e-06	-0.30	0.25		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)

1053

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1763	1fgv	A	231	450	1.3e-08	0.16	-0.05		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1763	1fgv	A	255	526	7e-09	-0.00	-0.20		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1763	1fgv	A	30	283	7.2e-18	0.41	-0.12		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1763	1fgv	A	71	373	2.4e-17	0.04	-0.07		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1763	1f88	A	250	313	5.6e-07	-0.25	0.04		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH REPEAT 2 (LRR) DOMAINS
1763	1iam		420	518	8.4e-16	0.10	0.04		INTERCELLULAR ADHESION MOLECULE-1; CHAIN: NUL1;	RHINOVIRUS RECEPTOR ICAM-1, CD54; RHINOVIRUS RECEPTOR, CELL ADHESION, INTEGRIN

1054

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1763	1hll	G	420	518	4.8e-15	0.30	0.27		HEPARIN-BINDING GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	LIGAND, 2 GLYCOPROTEIN, LFA-1 LIGAND, IMMUNOGLOBULIN FOLD, 3 TRANSMEMBRANE
1763	1itb	B	420	519	2.4e-13	0.17	-0.14		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
1763	1nct		409	501	1.1e-16	0.58	0.82		TITIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTIN5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
1763	1hnm		1	79	5.6e-18			53.79	MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1TNM4 1TNM58	
1763	1hnm		420	501	3.6e-16	0.55	1.00		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM3 (NMR, MINIMIZED AVERAGE STRUCTURE)	

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1763	1yr8	A	173	379	1.2e-13	0.26	0.25		1ITNM 4 1ITNM 58	
									GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1763	1yr8	A	65	289	9.6e-21	0.17	-0.02		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1763	1yr8	A	81	360	6e-24	0.42	0.31		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1763	2bnh		1	446	7.2e-36			83.20	RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1763	2bnh		60	405	4.2e-13	0.15	0.23		RIBONUCLEASE INHIBITOR; CHAIN:	ACETYLATION RNASE INHIBITOR,

1056

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									NULL;	RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLTATION, LEUCINE-RICH REPEATS
1763	2bnh		65	364	7.2e-36	0.33	0.75		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLTATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLTATION, LEUCINE-RICH REPEATS
1763	2fcb	A	411	518	2.4e-13	0.19	-0.01		FC GAMMA RIIB; CHAIN: A;	IMMUNE SYSTEM CD32; RECEPTOR, FC, CD32, IMMUNE SYSTEM
1763	3ncn	A	420	502	1.2e-15	0.30	0.92		NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM; CHAIN: A;	CELL ADHESION PROTEIN NCAM MODULE 2; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, HOMOPHILIC 3 BINDING, CELL ADHESION PROTEIN
1764	1atl	A	152	359	1.3e-28	0.81	1.00		ATROLYSIN C; 1ATL 4 CHAIN: A, B, C, D; 1ATL 5	METALLOENDOPETIDASE HEMORRHAGIC TOXIN C, FORM D; 1ATL 6
1764	1atl	A	156	359	1.2e-37	0.83	1.00		ATROLYSIN C; 1ATL 4 CHAIN: A, B, C, D; 1ATL 5	METALLOENDOPETIDASE HEMORRHAGIC TOXIN C, FORM D; 1ATL 6
1764	1bkc	A	156	357	3.6e-40	0.20	0.77		TUMOR NECROSIS FACTOR-ALPHA-CONVERTING ENZYME; CHAIN: A, C, E, I;	ZN-ENDOPEPTIDASE TACE; ZN-ENDOPEPTIDASE, HYDROLASE, TNF-ALPHA
1764	1bud	A	154	357	7e-27	0.68	1.00		ACUTOLYSIN A; CHAIN: A;	TOXIN HEMORRHAGIN I, 1AAH-1; METALLOPROTEINASE, SNAKE VENOM, MMP, TOXIN
1764	1dva	L	519	596	1.4e-12	0.10	-0.18		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN:	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE



1057

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									H, I, DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DEN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	COMPLEX
1764	1emn		126	194	7e-14	0.03	-0.20		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1764	1emn		327	394	1.4e-11	0.12	-0.20		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
1764	1hj7	A	358	438	4.2e-11	0.12	-0.20		LDL RECEPTOR; CHAIN: A;	CELL-SURFACE RECEPTOR CELL-SURFACE RECEPTOR, CALCIUM-BINDING, EGF-LIKE DOMAIN, 2 MODULE, APO-E, APO-B, LDL, VLDL
1764	1iag		152	359	2.8e-28	0.80	1.00		METALLOPROTEINASE ADAMALYSIN II (PROTEINASE II) (E.C.3.4.24.46) 1IAG 3	
1764	1iag		156	359	1.2e-37	0.84	1.00		METALLOPROTEINASE ADAMALYSIN II (PROTEINASE II) (E.C.3.4.24.46) 1IAG 3	

1058

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1764	1qua	A	152	357	5.6e-28	0.87	1.00		ACUTOLYSIN-C; CHAIN: A;	TOXIN HEMORRHAGIN III METALLOPROTEASE, HEMORRHAGIC TOXIN, SNAKE VENOM PROTEINASE, 2 CRYSTAL STRUCTURE, AGKISTRODON ACUTUS
1764	1qua	A	156	357	9.6e-32	0.77	1.00		ACUTOLYSIN-C; CHAIN: A;	TOXIN HEMORRHAGIN III METALLOPROTEASE, HEMORRHAGIC TOXIN, SNAKE VENOM PROTEINASE, 2 CRYSTAL STRUCTURE, AGKISTRODON ACUTUS
1764	9wga	A	400	568	8.4e-15	0.00	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
1768	1ewk	A	40	295	7e-42	-0.20	0.10		METABOTROPIC GLUTAMATE RECEPTOR SUBTYPE 1; CHAIN: A, B;	SIGNALING PROTEIN MGLUR1; SIGNAL TRANSDUCTION, NEUROTRANSMITTER, CNS, NEURON
1768	1ewk	A	42	223	9.6e-20	0.39	1.00		METABOTROPIC GLUTAMATE RECEPTOR SUBTYPE 1; CHAIN: A, B;	SIGNALING PROTEIN MGLUR1; SIGNAL TRANSDUCTION, NEUROTRANSMITTER, CNS, NEURON
1768	2lbp		48	174	9.6e-08	-0.06	0.06		PERIPLASMIC BINDING PROTEIN LEUCINE-BINDING PROTEIN (LBP5) 2LBP 4	
1768	2liv		48	142	0.00084	0.08	0.09		PERIPLASMIC BINDING PROTEIN LEUCINE(SLASH)*ISOLEUCINE(SLASH)*VALINE-BINDING PROTEIN 2LIV 4 (LIVBP5) 2LIV 5	

1059

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1768	2liv		48	221	3.6e-12	-0.06	0.21		PERIPLASMIC BINDING PROTEIN LEUCINE(SLASH)*ISOLEUCINE(SLASH)*VALINE-BINDING PROTEIN 2LIV 4 (LIVBP\$) 2LIV 5	
1769	1a25	A	2	109	5.6e-33			55.91	PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALB; CALCIUM+/+/PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN
1769	1a25	A	366	498	5.6e-30			77.89	PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALB; CALCIUM+/+/PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN
1769	1a25	A	369	495	5.6e-30	0.12	0.58		PROTEIN KINASE C (BETA); CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALB; CALCIUM+/+/PHOSPHOLIPID BINDING PROTEIN, 2 CALCIUM-BINDING PROTEIN
1769	1byn	A	1	109	1.4e-36			80.43	SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1769	1byn	A	1	109	7e-31			59.50	SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1769	1byn	A	231	356	1.4e-33	0.55	1.00		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS

1060

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1769	1byn	A	232	360	1.4e-33			95.19	SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1769	1byn	A	365	490	5.6e-23	0.21	1.00		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1769	1byn	A	367	493	9.6e-29	0.39	1.00		SYNAPTOTAGMIN I; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN, C2-DOMAIN, EXOCYTOSIS, NEUROTRANSMITTER 2 RELEASE, ENDOCYTOSIS/EXOCYTOSIS
1769	1dix	A	223	377	2.4e-21	0.33	0.13		PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C, CHAIN: A, B;	LIPID DEGRADATION PLC-DI; PHOSPHORIC DIESTER HYDROLASE, HYDROLASE, LIPID DEGRADATION, 2 TRANSDUCER, CALCIUM-BINDING, PHOSPHOLIPASE C, 3 PHOSPHOINOSITIDE-SPECIFIC
1769	1dix	B	228	377	3.6e-22	0.22	0.55		PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C, CHAIN: A, B;	LIPID DEGRADATION PLC-DI; PHOSPHORIC DIESTER HYDROLASE, HYDROLASE, LIPID DEGRADATION, 2 TRANSDUCER, CALCIUM-BINDING, PHOSPHOLIPASE C, 3

1061

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1769	1dqy	A	230	505	9.8e-65			235.90	SYNAPTOTAGMIN III; CHAIN: A;	PHOSPHOINOSITIDE-SPECIFIC ENDOCYTOSIS/EXOCYTOSIS
1769	1dqy	A	233	504	9.8e-65	0.24	1.00		SYNAPTOTAGMIN III; CHAIN: A;	BETA SANDWICH, CALCIUM ION, C2 DOMAIN
1769	1dsy	A	2	109	1.1e-33			57.02	PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS BETA SANDWICH, CALCIUM ION, C2 DOMAIN
1769	1dsy	A	366	504	4.2e-30			82.11	PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	TRANSPERASE CALCIUM <sup>++</sup> , PHOSPHOLIPID BINDING
1769	1dsy	A	369	495	4.2e-30	0.15	0.65		PROTEIN KINASE C, ALPHA TYPE; CHAIN: A;	PROTEIN, CALCIUM-BINDING 2
1769	1rlw		380	472	6e-20	0.07	-0.09		PHOSPHOLIPASE A2; CHAIN: NULL;	PROTEIN, CALCIUM-BINDING 2
1769	1rsy		1	109	1.4e-36			84.72	CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	PROTEIN, PHOSPHATIDYL SERINE, PROTEIN KINASE C
1769	1rsy		1	109	7e-31			61.17	CALCIUM/PHOSPHOLIPID	HYDROLASE CALB DOMAIN; HYDROLASE, C2 DOMAIN, CALB DOMAIN

1062

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									D BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
1769	1rsy		223	359	1.4e-33			103.60	CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
									CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
1769	1rsy		231	356	1.4e-33	0.47	1.00		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
									CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
1769	1rsy		363	490	2.8e-23	0.34	0.95		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
									CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
1769	1rsy		367	491	4.8e-28	0.50	1.00		CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
									CALCIUM/PHOSPHOLIPID BINDING PROTEIN SYNAPTOTAGMIN I (FIRST C2 DOMAIN) (CALB) IRSY 3	
1769	3rpb	A	367	502	2.8e-31	0.45	1.00		RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
									RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
1769	3rpb	A	367	504	1.2e-38	0.42	0.92		RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
									RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
1769	3rpb	A	367	505	1.2e-38			92.99	RABPHILIN 3-A; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS C2-DOMAINS, C2B-DOMAIN, RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS

1063

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1771	1edh	A	142	344	1.4e-49	0.09	0.95		E-CADHERIN; CHAIN: A, B;	RABPHILIN, ENDOCYTOSIS/EXOCYTOSIS
1771	1edh	A	23	127	1.3e-14	-0.15	0.25		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	251	452	1.2e-35	0.22	0.98		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	251	452	4.2e-29	-0.01	0.81		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	349	557	4.8e-46	0.27	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	35	236	1.2e-21	-0.22	0.04		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12;

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	360	539	1.4e-30	0.09	0.92		E-CADHERIN; CHAIN: A; B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	467	693	3.6e-27	0.12	0.47		E-CADHERIN; CHAIN: A; B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	494	697	7e-26	0.03	0.47		E-CADHERIN; CHAIN: A; B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1edh	A	66	236	1.4e-38	-0.34	0.39		E-CADHERIN; CHAIN: A; B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
1771	1ncg		142	235	9.8e-17	0.17	-0.08		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN CADHERIN INCG 13
1771	1ncg		251	343	8.4e-07	0.25	0.23		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN CADHERIN INCG 13
1771	1ncg		255	343	4.8e-13	-0.18	0.01		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN CADHERIN INCG 13
1771	1ncg		356	450	3.6e-06	0.05	0.51		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN



1065

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1771	Incg		360	450	7e-05	0.00	0.06		N-CADHERIN; INCG 3	CADHERIN INCG 13 CELL ADHESION PROTEIN
1771	Incg		467	556	3.6e-10	0.34	0.13		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN CADHERIN INCG 13
1771	Incg		491	542	2.8e-06	-0.05	0.63		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN CADHERIN INCG 13
1771	Incg		631	695	0.0017	0.06	0.10		N-CADHERIN; INCG 3	CELL ADHESION PROTEIN CADHERIN INCG 13
1771	Inci	B	142	236	2.8e-16	0.14	0.09		N-CADHERIN; INCI 3	CELL ADHESION PROTEIN CADHERIN INCI 13
1771	Inci	B	277	344	7e-07	0.09	0.13		N-CADHERIN; INCI 3	CELL ADHESION PROTEIN CADHERIN INCI 13
1771	Inci	B	295	344	9.6e-09	-0.28	0.66		N-CADHERIN; INCI 3	CELL ADHESION PROTEIN CADHERIN INCI 13
1771	Inci	B	398	452	2.4e-07	-0.47	0.53		N-CADHERIN; INCI 3	CELL ADHESION PROTEIN CADHERIN INCI 13
1771	Inci	B	495	542	5.6e-06	0.05	0.17		N-CADHERIN; INCI 3	CELL ADHESION PROTEIN CADHERIN INCI 13
1771	Incj	A	142	344	1.4e-52	0.17	0.83		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN ADHESION PROTEIN
1771	Incj	A	23	127	1.1e-15	-0.22	0.41		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN ADHESION PROTEIN
1771	Incj	A	251	452	5.6e-31	0.04	0.99		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN ADHESION PROTEIN
1771	Incj	A	356	560	2.4e-37	-0.02	0.24		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN ADHESION PROTEIN
1771	Incj	A	360	543	8.4e-31	0.17	0.96		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN ADHESION PROTEIN
1771	Incj	A	488	697	2.8e-27	0.19	0.25		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN ADHESION PROTEIN
1771	Incj	A	605	716	4.2e-10	-0.22	0.05		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN ADHESION PROTEIN
1771	Incj	A	66	236	8.4e-39	-0.45	0.36		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN

1066

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1771	1suh		251	348	2.8e-09	0.15	0.51		EPITHELIAL CADHERIN; CHAIN: NULL;	ADHESION PROTEIN
1771	1suh		253	348	2.4e-23	0.03	0.24		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1771	1suh		356	456	1.2e-15	0.06	0.17		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1771	1suh		360	456	8.4e-08	0.02	0.37		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1771	1suh		461	561	6e-24	0.39	0.46		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1771	1suh		494	561	5.6e-07	0.06	0.68		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1771	1suh		613	693	1.2e-09	0.28	-0.14		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
1778	1sfc	A	110	175	8.4e-25	-0.70	0.03		SYNAPTOBREVIN 2; CHAIN: A, E, J; SYNTAXIN 1A; CHAIN: B, F, J; SNAP-25B; CHAIN: C, G, K; SNAP-25B; CHAIN: D, H, L;	TRANSPORT PROTEIN VAMP 2; MEMBRANE FUSION PROTEIN COMPLEX, TRANSPORT PROTEIN
1780	1a17		76	242	5.6e-33			82.28	SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRAPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE

1067

Table 5

SEQ ID NO.	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1780	1aln	A	11	174	4.2e-74			125.43	B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1780	1aln	A	25	274	0	0.68	1.00		B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1780	1aln	A	25	274	0			324.34	B*3501; CHAIN: A, B; PEPTIDE VPLRPMTY; CHAIN: C;	COMPLEX (ANTIGEN/PEPTIDE) B35; MAJOR HISTOCOMPATIBILITY ANTIGEN, MHC, HLA, HLA-B3501, HIV, 2 NEF, COMPLEX (ANTIGEN/PEPTIDE)
1780	1agd	A	11	174	2.8e-75			121.80	B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL-INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1780	1agd	A	25	274	0	0.50	1.00		B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL-INDEX PEPTIDE); CHAIN: C;	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX
1780	1agd	A	25	274	0			322.00	B*0801; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HIV-1 GAG PEPTIDE (GGKKKYKL-INDEX PEPTIDE);	HISTOCOMPATIBILITY COMPLEX B8; B2M; PEPTIDE HLA B8, HIV, MHC CLASS I, HISTOCOMPATIBILITY COMPLEX

1068

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1780	1b42	E	197	283	1.2e-06	-0.26	0.90		CHAIN: C; HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)
1780	1ed3	A	11	174	1.4e-78			128.96	CLASS I MAJOR HISTOCOMPATIBILITY ANTIGEN RT1-AA; CHAIN: A, D; BETA-2- MICROGLOBULIN; CHAIN: B, E; PEPTIDE MTF-E (13N3E); CHAIN: C, F;	IMMUNE SYSTEM MAJOR HISTOCOMPATIBILITY COMPLEX, RAT MINOR 2 HISTOCOMPATIBILITY COMPLEX, MHC, IMMUNOLOGY, PEPTIDE 3 ANTIGEN PRESENTATION, CELLULAR IMMUNITY, CELL SURFACE 4 RECEPTOR, T CELL RECEPTOR LIGAND
1780	1ed3	A	25	273	0	0.44	1.00		CLASS I MAJOR HISTOCOMPATIBILITY ANTIGEN RT1-AA; CHAIN: A, D; BETA-2- MICROGLOBULIN; CHAIN: B, E; PEPTIDE MTF-E (13N3E); CHAIN: C, F;	IMMUNE SYSTEM MAJOR HISTOCOMPATIBILITY COMPLEX, RAT MINOR 2 HISTOCOMPATIBILITY COMPLEX, MHC, IMMUNOLOGY, PEPTIDE 3 ANTIGEN PRESENTATION, CELLULAR IMMUNITY, CELL SURFACE 4 RECEPTOR, T CELL RECEPTOR LIGAND
1780	1ed3	A	25	273	0			306.45	CLASS I MAJOR HISTOCOMPATIBILITY ANTIGEN RT1-AA; CHAIN: A, D; BETA-2- MICROGLOBULIN; CHAIN: B, E; PEPTIDE MTF-E (13N3E); CHAIN: C, F;	IMMUNE SYSTEM MAJOR HISTOCOMPATIBILITY COMPLEX, RAT MINOR 2 HISTOCOMPATIBILITY COMPLEX, MHC, IMMUNOLOGY, PEPTIDE 3 ANTIGEN PRESENTATION, CELLULAR IMMUNITY, CELL

1069

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	PsI Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1780	1efx	A	11	174	5.6e-76			129.79	C, F;	SURFACE 4 RECEPTOR, T CELL RECEPTOR LIGAND
									HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1780	1efx	A	25	276	0	0.51	1.00		HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1780	1efx	A	25	276	0			332.02	HLA-CW3 (HEAVY CHAIN); CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; PEPTIDE FROM IMPORTIN ALPHA-2; CHAIN: C; NATURAL KILLER CELL RECEPTOR KIR2DL2; CHAIN: D, E;	IMMUNE SYSTEM MHC, HLA, CLASS I, KIR, NK CELL RECEPTOR, IMMUNOGLOBULIN 2 FOLD, RECEPTOR/MHC COMPLEX
1780	1efr	A	89	217	4.2e-22			52.69	TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL -D score	Compound	PDB annotation
1780	1elw	A	86	204	1.3e-26			61.34	CHAIN: B; TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70- PEPTIDE; CHAIN: C, D;	BINDING CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
1780	1f3o	A	1	199	1.4e-48			51.22	HYPOTHETICAL ABC TRANSPORTER ATP- BINDING PROTEIN CHAIN: A;	STRUCTURAL GENOMICS TRANSPORTER
1780	1fzk	A	11	174	8.4e-75			118.23	H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, K-B CHAIN: A; BETA-2- MICROGLOBULIN; CHAIN: B; NUCLEOCAPSID PROTEIN; CHAIN: P;	IMMUNE SYSTEM SEV9; MAJOR HISTOCOMPATIBILITY COMPLEX PEPTIDE-MHC
1780	1fzk	A	25	272	0	0.59	1.00		H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, K-B CHAIN: A; BETA-2- MICROGLOBULIN; CHAIN: B; NUCLEOCAPSID PROTEIN; CHAIN: P;	IMMUNE SYSTEM SEV9; MAJOR HISTOCOMPATIBILITY COMPLEX PEPTIDE-MHC
1780	1fzk	A	25	272	0			293.35	H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, K-B CHAIN: A; BETA-2- MICROGLOBULIN; CHAIN: B; NUCLEOCAPSID PROTEIN; CHAIN: P;	IMMUNE SYSTEM SEV9; MAJOR HISTOCOMPATIBILITY COMPLEX PEPTIDE-MHC
1780	1hoc	A	11	174	5.6e-74			121.28	HISTOCOMPATIBILITY ANTIGEN MURINE	

1071

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX CONSISTING 1HOC 3 OF H-2D=B=, B2-MICROGLOBULIN, AND A 9-RESIDUE PEPTIDE 1HOC 4	
1780	1hoc	A	25	270	0			293.21	HISTOCOMPATIBILITY ANTIGEN MURINE CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX CONSISTING 1HOC 3 OF H-2D=B=, B2-MICROGLOBULIN, AND A 9-RESIDUE PEPTIDE 1HOC 4	
1780	1hsa	A	11	174	5.6e-74			125.78	HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN 1HSA 3 /HLA-B(ASTERISK)2705\$ 1HSA 4	
1780	1hsa	A	25	274	0	0.67	1.00		HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN 1HSA 3 /HLA-B(ASTERISK)2705\$ 1HSA 4	
1780	1hsa	A	25	274	0			324.99	HISTOCOMPATIBILITY ANTIGEN HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN 1HSA 3 /HLA-	

1072

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									B(ASTERISK)2705\$ IHS A <sup>4</sup>	
1780	1hsb	A	11	174	1.4e-75			118.41	HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE IHSB 3 ANTIGEN) IHSB 4	
1780	1hsb	A	25	268	0			299.60	HISTOCOMPATIBILITY ANTIGEN CLASS I HISTOCOMPATIBILITY ANTIGEN AW68.1 (LEUCOCYTE IHSB 3 ANTIGEN) IHSB 4	
1780	1i4f	A	11	174	2.8e-74			129.54	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1780	1i4f	A	25	273	0	0.55	1.00		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C;	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED ANTIGEN
1780	1i4f	A	25	273	0			324.98	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 CHAIN: A; BETA-2-MICROGLOBULIN;	IMMUNE SYSTEM MAGE-4 ANTIGEN; MAJOR HISTOCOMPATIBILITY COMPLEX, HUMAN LEUKOCYTE ANTIGEN, 2 MELANOMA-ASSOCIATED



1073

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1780	1iam		182	282	8.4e-07	0.08	-0.03		CHAIN: B; MELANOMA-ASSOCIATED ANTIGEN 4; CHAIN: C; INTERCELLULAR ADHESION MOLECULE-1; CHAIN: NULL;	ANTIGEN RHINOVIRUS RECEPTOR ICAM-1, CD54; RHINOVIRUS RECEPTOR, CELL ADHESION, INTEGRIN LIGAND, 2 GLYCOPROTEIN, LFA-1 LIGAND, IMMUNOGLOBULIN FOLD, 3 TRANSMEMBRANE
1780	1ld9	A	11	174	1.3e-75			125.09	MHC CLASS I H-2LD HEAVY CHAIN; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; NANO-PEPTIDE; CHAIN: C;	MAJOR HISTOCOMPATIBILITY COMPLEX LD; MAJOR HISTOCOMPATIBILITY COMPLEX, LD
1780	1ld9	A	25	266	0			287.52	MHC CLASS I H-2LD HEAVY CHAIN; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; NANO-PEPTIDE; CHAIN: C;	MAJOR HISTOCOMPATIBILITY COMPLEX LD; MAJOR HISTOCOMPATIBILITY COMPLEX, LD
1780	1mco	H	193	283	0.0072	0.44	0.30		IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION IMCO 3	
1780	1mhc	A	11	174	1.4e-62			92.09	MHC CLASS I ANTIGEN H2-M3; 1MHC 6 CHAIN: A, B, D, E; 1MHC 7 NONAPEPTIDE FROM RAT NADH DEHYDROGENASE; 1MHC 12 CHAIN: C, F; 1MHC 13	HISTOCOMPATIBILITY ANTIGEN/PEPTIDE MAJOR HISTOCOMPATIBILITY COMPLEX; 1MHC 8 NDI; 1MHC 15
1780	1mhc	A	25	274	2.8e-98			274.19	MHC CLASS I ANTIGEN	HISTOCOMPATIBILITY

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									H2-M3; 1MHC 6 CHAIN: A, B, D, E; 1MHC 7 NONAPEPTIDE FROM RAT NADH DEHYDROGENASE; 1MHC 12 CHAIN: C, F; 1MHC 13	ANTIGEN/PEPTIDE MAJOR HISTOCOMPATIBILITY COMPLEX; 1MHC 8 ND1; 1MHC 15
1780	1mhe	A	12	174	4.2e-72			124.46	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-E; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D; PEPTIDE (VMAPRTVL); CHAIN: P, Q;	MAJOR HISTOCOMPATIBILITY COMPLEX MHC NONCLASSICAL CHAIN, MHC-E, HLA-E, MHC CLASS HLA-E, HLA E, MAJOR HISTOCOMPATIBILITY COMPLEX, MHC, HLA, 2 BETA 2 MICROGLOBULIN, PEPTIDE, LEADER PEPTIDE, 3 NON-CLASSICAL MHC, CLASS IB MHC
1780	1mhe	A	26	272	0	0.34	1.00		HLA CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-E; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D; PEPTIDE (VMAPRTVL); CHAIN: P, Q;	MAJOR HISTOCOMPATIBILITY COMPLEX MHC NONCLASSICAL CHAIN, MHC-E, HLA-E, MHC CLASS HLA-E, HLA E, MAJOR HISTOCOMPATIBILITY COMPLEX, MHC, HLA, 2 BETA 2 MICROGLOBULIN, PEPTIDE, LEADER PEPTIDE, 3 NON-CLASSICAL MHC, CLASS IB MHC
1780	1mhe	A	26	272	0			321.02	HLA CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-E; CHAIN: A, C; BETA-2-MICROGLOBULIN; CHAIN: B, D; PEPTIDE (VMAPRTVL); CHAIN: P, Q;	MAJOR HISTOCOMPATIBILITY COMPLEX MHC NONCLASSICAL CHAIN, MHC-E, HLA-E, MHC CLASS HLA-E, HLA E, MAJOR HISTOCOMPATIBILITY COMPLEX, MHC, HLA, 2 BETA 2 MICROGLOBULIN, PEPTIDE, LEADER PEPTIDE, 3 NON-CLASSICAL MHC, CLASS IB MHC
1780	1qo3	A	12	174	1.4e-74			120.50	MHC CLASS I H-2DD	COMPLEX (NK RECEPTOR/MHC

1075

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C; D;	CLASS D H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
1780	1qo3	A	26	272	0	0.60	1.00		MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C; D;	COMPLEX (NK RECEPTOR/MHC CLASS D) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
1780	1qo3	A	26	273	0			299.18	MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C; D;	COMPLEX (NK RECEPTOR/MHC CLASS D) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
1780	1qqd	A	12	174	2.8e-74			127.23	HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1780	1qqd	A	26	272	0	0.55	1.00		HISTOCOMPATIBILITY	IMMUNE SYSTEM

1076

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1780	1qgd	A	26	272	0			321.09	HISTOCOMPATIBILITY LEUKOCYTE ANTIGEN (HLA)-CW4 CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; HLA-CW4 SPECIFIC PEPTIDE; CHAIN: C;	IMMUNE SYSTEM IMMUNOGLOBULIN (IG)-LIKE DOMAIN, ALPHA HELIX, BETA SHEET, 2 IMMUNE SYSTEM
1780	1tmc	A	11	174	1.4e-75			164.71	HISTOCOMPATIBILITY ANTIGEN TRUNCATED HUMAN CLASS I HISTOCOMPATIBILITY ANTIGEN HLA-AW68 1TMC 3 COMPLEXED WITH A DECAMERIC PEPTIDE (EVAPPEYHRK) 1TMC 4	
1782	1b0u	A	426	619	7e-41	0.52	1.00		HISTIDINE PERMEASE; CHAIN: A;	TRANSPORT PROTEIN ABC TRANSPORTER, HISP, ABC TRANSPORTER, HISTIDINE PERMEASE, TRANSPORT PROTEIN
1782	1cex		624	775	6e-09	0.61	-0.19		CUTTNASE; CHAIN: NULL;	SERINE ESTERASE HYDROLASE, SERINE ESTERASE, GLYCOPROTEIN
1782	1d6j	A	446	536	0.0041	0.00	0.05		ADENOSINE-5'PHOSPHOSULFATE KINASE; CHAIN: A, B;	TRANSFERASE APS KINASE; APS KINASE, ADENYLYLSULFATE KINASE, SULFATE, NUCLEOTIDE

1077

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1782	1e69	A	447	616	2.4e-07	-0.04	0.04		CHROMOSOME SEGREGATION SMC PROTEIN; CHAIN: A, B, C, D, E, F;	2 KINASE, TRANSFERASE
1782	1ex7	A	451	477	0.0055	-0.66	0.24		GUANYLATE KINASE; CHAIN: A;	TRANSFERASE GUANYLATE KINASE, SUBSTRATE-INDUCED FIT, DOMAIN MOVEMENT, 2 GMP, ATP, SUBSTRATE SPECIFICITY, X-RAY DIFFRACTION
1782	1f3o	A	421	619	2.8e-48	0.52	1.00		HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CHAIN: A;	STRUCTURAL GENOMICS TRANSPORTER
1782	1f3o	A	438	620	6e-22	0.55	1.00		HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CHAIN: A;	STRUCTURAL GENOMICS TRANSPORTER
1782	1fuu	B	445	599	0.0096	-0.08	0.10		YEAST INITIATION FACTOR 4A; CHAIN: A, B;	TRANSLATION EUKARYOTIC INITIATION FACTOR 4A; IF4A, HELICASE, DEAD-BOX PROTEIN
1782	1g29	1	423	614	9.8e-49	0.36	1.00		MALTOSE TRANSPORT PROTEIN MALK; CHAIN: 1, 2;	SUGAR BINDING PROTEIN MALK; ATPASE, ACTIVE TRANSPORT, MALTOSE UPTAKE AND REGULATION
1782	1g6h	A	419	614	2.4e-20	0.44	1.00		HIGH-AFFINITY BRANCHED-CHAIN AMINO ACID CHAIN: A	TRANSPORT PROTEIN BETA-CORE DOMAIN, ABC SPECIFIC-BETA-STRAND DOMAIN ALPHA-2 HELIX DOMAIN
1782	1g6h	A	432	619	2.8e-30	0.47	0.69		HIGH-AFFINITY BRANCHED-CHAIN AMINO ACID CHAIN: A	TRANSPORT PROTEIN BETA-CORE DOMAIN, ABC SPECIFIC-BETA-STRAND DOMAIN ALPHA-2 HELIX DOMAIN
1782	1g6o	A	412	536	1.2e-08	-0.08	0.00		CAG-ALPHA; CHAIN: A;	HYDROLASE TRAFFIC ATPASE;

1078

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1782	1ga6	A	624	763	9.6e-10	0.82	-0.19		B; SERINE-CARBOXYL PROTEINASE; CHAIN: A; FRAGMENT OF TYROSTATIN; CHAIN: I;	ATPASE, TYPE IV SECRETION SYSTEM
1782	1ga6	A	632	773	3.6e-09	0.65	-0.18		SERINE-CARBOXYL PROTEINASE; CHAIN: A; FRAGMENT OF TYROSTATIN; CHAIN: I;	HYDROLASE PSCP, PSEUDOMONAPERSIN, PEPSTATIN-INSENSITIVE SERINE- CARBOXYL PROTEINASE
1782	1hey		450	543	0.0031	0.10	0.62		CHEMOTAXIS CHEY MUTANT WITH ASP 12 REPLACED BY GLY, ASP 13 REPLACED BY IHEY 3 ASN, PHE 14 REPLACED BY GLY, SER 15 REPLACED BY GLY, MET 17 IHEY 4 REPLACED BY GLY, ARG 18 REPLACED BY LYS, ARG 19 REPLACED BY IHEY 5 SER, ILE 20 REPLACED BY THR, GLU 35 REPLACED BY ASP (D12G, IHEY 6 D13N, F14G, S15G, M17G, R 18K, R19S, I20T, E35D) (SYNCHROTRON X-RAY IHEY 7 DIFFRACTION) IHEY 8	
1782	1kap	P	623	772	1.2e-06	1.06	-0.19		ALKALINE PROTEASE; IKAP 4 CHAIN: P; IKAP 5 TETRAPEPTIDE (GLY SER ASN SER); IKAP 9 CHAIN: I; IKAP 10	ZINC METALLOPROTEASE P. AERUGINOSA ALKALINE PROTEASE; IKAP 6 CALCIUM BINDING PROTEIN IKAP 19

1079

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1782	1qg4	A	607	760	1.2e-08	0.83	-0.20		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1782	1qg4	A	668	774	2.4e-09	0.98	-0.19		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
1782	1tal		627	774	4.8e-09	0.81	-0.19		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
1782	2por		626	760	3.6e-07	0.83	-0.19		INTEGRAL MEMBRANE PROTEIN PORIN PORIN (CRYSTAL FORM B) 2POR 3	
1782	3pm		625	780	1.2e-06	0.88	-0.19		PORIN; CHAIN: NULL;	MEMBRANE PROTEIN INTEGRAL MEMBRANE PROTEIN, PORIN, PORE EYELET MUTANT
1784	1qho	A	1022	1129	0.0039	0.39	0.11		ALPHA-AMYLASE; CHAIN: A;	HYDROLASE "MALTOGENIC" ALPHA AMYLASE; AMYLASE, GLYCOSIDE HYDROLASE, STARCH DEGRADATION
1784	1wer		1885	2041	3.6e-32	-0.23	0.07		P120GAP; CHAIN: NULL;	GTPASE ACTIVATION GAP-334, GAPETTE; GTPASE ACTIVATION, RAS, GAP, SIGNAL TRANSDUCTION, GROWTH 2 REGULATION, CANCER
1786	1a4y	A	192	638	8.4e-42	0.23	0.90		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS

1080

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1786	1a4y	A	324	677	6e-32	0.07	0.42		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RT-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1786	1a4y	A	350	677	4.2e-20	-0.18	0.30		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RT-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1786	1a4y	A	407	730	7e-17	-0.38	0.09		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RT-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPTOPE MAPPING, LEUCINE-RICH 3 REPEATS
1786	1a9n	A	133	278	1.2e-15	-0.37	0.19		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	A	149	312	3.6e-18	0.03	-0.12		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	A	385	495	1.2e-19	-0.17	0.72		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'';	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA,



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1786	1a9n	A	404	575	4.8e-23	0.33	0.81		CHAIN: B, D; U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	SNRNP, RIBONUCLEOPROTEIN COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	A	438	614	8.4e-18	0.02	-0.09		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	A	524	676	2.4e-20	0.15	0.62		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	C	130	276	9.6e-15	-0.12	0.09		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	C	384	490	2.4e-19	0.15	0.82		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	C	404	583	2.4e-24	0.33	0.78		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1a9n	C	524	676	1.2e-20	0.06	0.47		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B''; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1786	1d0b	A	132	332	4.8e-16	0.39	0.90		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	189	442	2.4e-18	0.29	0.34		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING,

1082

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1786	1d0b	A	211	391	1.4e-21	-0.15	0.83		INTERNALIN B; CHAIN: A;	CELL ADHESION
1786	1d0b	A	239	490	3.6e-21	0.09	0.76		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	248	441	1.4e-21	0.21	0.22		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	332	505	2.8e-25	0.44	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	369	554	1.2e-22	-0.07	0.06		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	36	273	4.2e-19	0.16	0.03		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	428	605	2.8e-22	0.06	0.78		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	461	676	7.2e-25	0.14	0.69		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	494	651	7e-24	0.28	0.96		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	521	678	1.4e-23	0.13	0.17		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1786	1d0b	A	566	698	8.4e-18	0.08	0.25		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION

1083

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1786	Idce	A	187	278	1.3e-09	0.59	0.37		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1786	Idce	A	354	444	1.4e-08	0.22	0.05		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1786	Idce	A	435	556	9.8e-10	0.13	0.37		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1786	Idce	A	483	581	1.1e-10	-0.23	0.25		RAB GERANYLGERANYLTR NSFERASE ALPHA SUBUNIT; CHAIN: A, C, RAB GERANYLGERANYLTR NSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS E, 2.0 A 2 RESOLUTION, N- FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1786	Idce	A	572	678	9.8e-12	-0.08	0.22		RAB GERANYLGERANYLTR NSFERASE ALPHA	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTR NSFERAS

1084

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	E, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1786	1dce	A	625	729	7e-08	0.02	0.36		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1786	1ds9	A	141	278	2.4e-18	-0.25	0.18		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1786	1ds9	A	257	414	1.4e-08	-0.29	0.40		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1786	1ds9	A	354	486	2.8e-11	-0.09	0.29		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1786	1ds9	A	404	537	1.2e-19	-0.52	0.98		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1786	1ds9	A	465	650	8.4e-15	-0.51	0.05		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA

1085

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1786	1ds9	A	556	677	8.4e-13	-0.28	0.11		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1786	1ds9	A	581	702	2.8e-11	-0.21	0.49		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1786	1foi	A	454	511	1.4e-06	-0.61	0.18		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1), RIBONUCLEOPROTEIN (RNP, RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
1786	1fqv	A	133	274	2.4e-10	0.12	0.49		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1786	1fqv	A	167	443	7.2e-11	0.06	0.15		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1786	1fqv	A	244	481	5.6e-14	0.18	-0.01		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX,

1086

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1786	1f9v	A	361	623	4.2e-14	0.12	0.71		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1786	1f9v	A	5	216	4.2e-09	0.12	-0.19		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1786	1f88	A	454	511	1.4e-06	-0.38	0.07		TIP ASSOCIATING PROTEIN; CHAIN: A, B, C, D, E;	RNA BINDING PROTEIN TAP; RIBONUCLEOPROTEIN (RNP, RRM, RBD) AND LEUCINE-RICH-REPEAT 2 (LRR) DOMAINS
1786	1f9v	A	777	932	1.4e-53	0.46	1.00		TOLL-LIKE RECEPTOR 1; CHAIN: A;	SIGNALING PROTEIN BETA-ALPHA-BETA FOLD PARALLEL BETA SHEET
1786	1f9x	A	789	932	5.6e-44	0.40	1.00		TOLL-LIKE RECEPTOR 2; CHAIN: A;	SIGNALING PROTEIN BETA-ALPHA-BETA FOLD
1786	1y85	A	113	390	1.4e-10	0.17	-0.12		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY

1087

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1786	1yrg	A	402	677	2.4e-29	-0.05	0.22		GTPASE-ACTIVATING PROTEIN RNAI_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1786	2bnh		134	493	3.6e-31	-0.11	0.09		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1786	2bnh		246	656	8.4e-37	-0.00	0.03		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1786	2bnh		391	730	1.1e-17	-0.15	0.57		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1786	2bnh		8	431	1.4e-19	0.12	-0.03		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
1792	1auk		554	657	3.6e-15	-0.19	0.17		ARYLSULFATASE A; CHAIN: NULL;	HYDROLASE CEREBROSIDE-3-SULFATE-SULFATASE; CEREBROSIDE-3-SULFATE HYDROLYSIS, LYSOSOMAL ENZYME, 2 HYDROLASE
1792	1dpi		837	921	9.8e-35	-0.32	0.93		NUCLEOTIDYLTRANSFER	

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1792	1fsu		548	691	6e-13	-0.22	0.19		N- RASE/DNA\$ POLYMERASE I (KLENOW FRAGMENT) (E.C.2.7.7.7) - \$D/CMP\$ 1DPI 4 COMPLEX 1DPI 5	HYDROLASE ARYLSULFATASE B, ASB, 4-SULFATASE, SULFATASE, GLYCOSAMINOGLYCAN DEGRADATION, HYDROLASE, SIGNAL, 2 GLYCOPROTEIN, LYSOSOME
1792	1fsu		591	630	0.0013	-0.02	0.22		N- ACETYL GALACTOSAMINE-4-SULFATASE; CHAIN: NULL;	HYDROLASE ARYLSULFATASE B, ASB, 4-SULFATASE, SULFATASE, GLYCOSAMINOGLYCAN DEGRADATION, HYDROLASE, SIGNAL, 2 GLYCOPROTEIN, LYSOSOME
1792	2kfn	A	837	921	9.8e-35	0.25	1.00		KLENOW FRAGMENT; CHAIN: A; DEOXYRIBO-3-S- PHOSPHOROTHIOLATE DNA; CHAIN: B;	COMPLEX (POLYMERASE/DNA) LARGE FRAGMENT, COMPLEX (POLYMERASE/DNA), EXONUCLEASE
1796	1a9n	A	284	415	2.4e-16	0.18	0.23		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1796	1a9n	C	284	402	8.4e-16	0.25	-0.07		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B'; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP, RIBONUCLEOPROTEIN
1796	1d0b	A	128	320	8.4e-19	0.12	0.01		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1796	1d0b	A	200	368	4.8e-11	0.40	1.00		INTERNALIN B; CHAIN:	CELL ADHESION LEUCINE RICH



Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A;	REPEAT, CALCIUM BINDING, CELL ADHESION
1796	1d0b	A	244	417	9.8e-21	0.19	0.62		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1796	1d0b	A	309	466	1.4e-23	0.13	0.12		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
1796	1dce	A	200	383	3.6e-10	0.13	-0.18		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1796	1dce	A	292	399	1.1e-10	0.50	0.98		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
1796	1ds9	A	295	368	3.6e-11	-0.57	0.13		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
1796	1ev2	B	437	528	4.8e-07	0.28	0.96		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGFR2; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD

1090

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1796	1fkg	A	431	526	4.8e-06	0.38	0.17		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
1796	1fgv	A	225	458	4.2e-09	-0.18	0.00		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
1796	1yrg	A	200	382	6e-10	0.14	-0.18		GTPASE-ACTIVATING PROTEIN RNAI_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1796	1yrg	A	298	392	1.2e-07	-0.17	0.04		GTPASE-ACTIVATING PROTEIN RNAI_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1796	1yrg	A	314	580	2.8e-12	0.21	-0.15		GTPASE-ACTIVATING PROTEIN RNAI_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNAIP; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPII, GTPASE-ACTIVATING PROTEIN, GAP, RNAIP, RANGAP, LRR, LEUCINE-2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL

1091

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1796	1yrg	A	78	350	2.8e-10	0.03	-0.18		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
1800	1ael	A	175	345	1.4e-10	-0.12	0.09		TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
1800	1ael	B	175	347	8.4e-11	-0.11	0.03		TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
1800	1edo	A	179	287	5.6e-07	0.16	0.10		BETA-KETO ACYL CARRIER PROTEIN REDUCTASE; CHAIN: A;	OXIDOREDUCTASE NUCLEOTIDE FOLD, ROSSMANN FOLD
1800	1fjh	A	180	281	2.8e-06	0.12	0.78		3ALPHA-HYDROXYSTEROID DEHYDROGENASE/CAR BONYL CHAIN: A, B;	OXIDOREDUCTASE 3ALPHA-HSD/CR, HYDROXYSTEROID SHORT CHAIN SHORT CHAIN DEHYDROGENASE, SDR, CARBONYL REDUCTASE, STEROID, 2 HYDROXYSTEROID, XENOBIOTIC, METYRAPONE, OLIGOMERISATION, 3 COMAMONAS TESTOSTERONI

1092

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1800	1fnc	A	175	349	2.8e-10	-0.22	0.00		7 ALPHA-HYDROXYSTEROID DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE/REDUCTASE, BILE ACID CATABOLISM
1800	1i01	A	174	279	1.4e-11	0.06	0.60		BETA-KETOACYL [ACP] REDUCTASE; CHAIN: A, B, C, D, E, F, G, H;	OXIDOREDUCTASE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE, 3-ROSSMAN FOLD
1800	1i01	C	174	279	5.6e-11	0.03	0.41		BETA-KETOACYL [ACP] REDUCTASE; CHAIN: A, B, C, D, E, F, G, H;	OXIDOREDUCTASE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE, 3-ROSSMAN FOLD
1800	1jb1	A	169	212	0.00098	-0.09	0.45		HPRK PROTEIN; CHAIN: A;	TRANSFERASE/HYDROLASE CATABOLITE REPRESSION, HPR PHOSPHORYLATION, LACTOBACILLUS 2 CASEI, P-LOOP, PROTEIN KINASE, HEXAMER
1805	1fnn	A	64	199	0.0024	0.38	0.09		CELL DIVISION CONTROL PROTEIN 6; CHAIN: A, B;	CELL CYCLE CDC6P, CDC6, CDC18, ORC1, AAA PROTEIN, DNA REPLICATION INITIATION 2 FACTOR, CELL CYCLE CONTROL FACTOR
1805	1g41	A	67	290	8.4e-11	-0.09	0.18		HEAT SHOCK PROTEIN HSLU; CHAIN: A;	CHAPERONE AAA-ATPASE, CLPY, ATP-DEPENDENT PROTEOLYSIS
1805	1im2	A	67	290	2.8e-09	0.13	-0.08		ATP-DEPENDENT HSL PROTEASE ATP-BINDING SUBUNIT CHAIN: A;	CHAPERONE HEAT SHOCK PROTEIN HSLU; CHAPERONE, AAA FAMILY
1809	1ck4	A	126	202	5.9e-05	0.08	-0.20		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
1809	1ido		125	202	0.003	0.14	-0.20		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN,

1093

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1809	1qcs	A	125	202	0.00012	0.40	-0.20		ALPHA1 BETA1 INTEGRIN; CHAIN: A; ALPHA1 BETA1 INTEGRIN; CHAIN: B;	EXTRACELLULAR 2 MATRIX, CYTOSKELETON
1813	1alh	A	412	507	7.5e-24	0.04	-0.20		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1813	1alh	A	438	535	1.3e-28	0.07	-0.20		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1813	1alh	A	739	818	3e-23	0.13	-0.20		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
1813	1mey	C	267	332	3e-35	0.43	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1813	1mey	C	383	463	6e-42	0.06	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

1094

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1813	Imey	C	414	508	1.8e-11	0.00	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1813	Imey	C	768	846	1.5e-39	0.29	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1813	Imey	C	796	852	3e-24	0.14	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1813	Imey	G	766	793	1.2e-12	0.11	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
1813	Iubd	C	741	846	9e-28	0.10	-0.20		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
1813	2adr		438	509	4.5e-13	0.15	-0.20		ADRI; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADRI, ZINC FINGER, NMR
1813	2gli	A	173	319	1.5e-28	0.00	-0.20		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

1095

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1813	2gli	A	445	594	3e-34	0.05	-0.20		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
1814	1f0h	A	5	153	0.0018	0.11	-0.20		PHENOL HYDROXYLASE; CHAIN: A, B, C, D;	FLAVIN FLAVIN, PHENOL HYDROXYLASE, MONOOXYGENASE, OXIDOREDUCTASE
1815	1a0j	A	123	334	4.5e-80			77.31	TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
1815	1a0j	A	140	362	7.5e-86			87.11	TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
1815	1a0l	A	141	332	1.5e-77			103.34	BETA-TRYPTASE; CHAIN: A, B, C, D;	SERINE PROTEINASE TRYPSIN- LIKE SERINE PROTEINASE, TETRAMER, HEPARIN, ALLERGY, 2 ASTHMA
1815	1a0l	A	158	380	6e-85			111.98	BETA-TRYPTASE; CHAIN: A, B, C, D;	SERINE PROTEINASE TRYPSIN- LIKE SERINE PROTEINASE, TETRAMER, HEPARIN, ALLERGY, 2 ASTHMA
1815	1a5e		9	182	9e-25			64.19	TUMOR SUPPRESSOR P16INK4A; CHAIN: NULL;	ANTI-ONCOGENE CELL CYCLE, ANTI-ONCOGENE, REPEAT, ANK REPEAT
1815	1a5i	A	96	334	4.5e-61			76.22	PLASMINOGEN ACTIVATOR; CHAIN: A; GLU-GLY-ARG CHLOROMETHYL KETONE; CHAIN: I;	COMPLEX (SERINE PROTEASE/INHIBITOR) (DELTA FEK) DSPAALPHA1; EGRCMK; SERINE PROTEASE, FIBRINOLYTIC ENZYMES, PLASMINOGEN 2 ACTIVATORS
1815	1aut	C	137	334	1.5e-58			80.06	ACTIVATED PROTEIN C; CHAIN: C, L, D-PHE-PRO- MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA;

1096

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
										HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
1815	1aut	C	154	365	1e-70			89.76	ACTIVATED PROTEIN C, CHAIN: C, L; D-PHE-PRO- MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
1815	1awc	B	22	168	7.5e-40			71.90	GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA- BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
1815	1bd8		1	154	9e-30			71.86	P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
1815	1bi7	B	22	152	6e-25			61.32	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; MULTIPLE TUMOR SUPPRESSOR; CHAIN: B;	COMPLEX (KINASE/ANTI- ONCOGENE) CDK6, P16INK4A, MTS1; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, MULTIPLE TUMOR SUPPRESSOR, 3 MTS1, COMPLEX (KINASE/ANTI- ONCOGENE) HEADER
1815	1bio		140	379	7.5e-61			85.15	COMPLEMENT FACTOR	SERINE PROTEASE SERINE



1097

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									D; CHAIN: NULL;	PROTEASE, HYDROLASE, COMPLEMENT, FACTOR D, CATALYTIC 2 TRIAD, SELF-REGULATION
1815	1blx	B	2	153	3e-29			64.53	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
1815	1bru	P	116	334	9e-71			80.45	ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
1815	1bru	P	133	380	3e-79			93.17	ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
1815	1bu9	A	20	200	1.5e-35			66.00	CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN-2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
1815	1c5y	B	139	362	1.5e-71			85.63	UROKINASE-TYPE PLASMINOGEN ACTIVATOR; CHAIN: A; UROKINASE-TYPE PLASMINOGEN ACTIVATOR; CHAIN: B;	BLOOD CLOTTING SELECTIVE, SITE INHIBITOR, STRUCTURE-BASED DRUG DESIGN, 2 UROKINASE, TRYPSIN, THROMBIN
1815	1cgh	A	120	334	6e-55			82.82	CATHEPSIN G; CHAIN: A; PHOSPHONATE INHIBITOR SUC-VAL-PRO-PHEP-(OPH)2; CHAIN: S;	COMPLEX (SERINE PROTEASE/INHIBITOR) INFLAMMATION, INHIBITOR, SPECIFICITY, SERINE PROTEASE, 2 COMPLEX (SERINE PROTEASE/INHIBITOR)
1815	1cgh	A	137	361	3e-54			88.26	CATHEPSIN G; CHAIN:	COMPLEX (SERINE

1098

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									A; PHOSPHONATE INHIBITOR SUC-VAL-PRO-PHEP-(OPH) <sub>2</sub> ; CHAIN: S;	PROTEASE/INHIBITOR) INFLAMMATION, INHIBITOR, SPECIFICITY, SERINE PROTEASE, 2 COMPLEX (SERINE PROTEASE/INHIBITOR)
1815	1chg		102	334	1.5e-66			83.30	HYDROLASE ZYMOGEN (SERINE PROTEINASE) CHYMOTRYPSINOGEN A 1CHG 4	
1815	1chg		119	376	4.5e-70			92.78	HYDROLASE ZYMOGEN (SERINE PROTEINASE) CHYMOTRYPSINOGEN A 1CHG 4	
1815	1d6w	A	78	334	6e-64			76.95	THROMBIN; CHAIN: A; DECAPEPTIDE INHIBITOR; CHAIN: I;	HYDROLASE/HYDROLASE INHIBITOR HYDROLASE, THROMBIN, THROMBIN INHIBITOR
1815	1dan	H	123	334	6e-64			82.55	BLOOD COAGULATION FACTOR VIIA; CHAIN: L; H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETO NE (DFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
1815	1dan	H	140	383	3e-73			101.34	BLOOD COAGULATION FACTOR VIIA; CHAIN: L; H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETO NE (DFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
1815	1ddj	A	106	334	7.5e-74			89.66	PLASMINOGEN; CHAIN: A, B, C, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC

1099

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1815	1ddj	A	123	362	9e-81			97.69	PLASMINOGEN; CHAIN: A, B, C, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC DOMAIN
1815	1ddj	A	125	380	9e-81	0.07	-0.20		PLASMINOGEN; CHAIN: A, B, C, D;	BLOOD CLOTTING PLASMINOGEN, CATALYTIC DOMAIN
1815	1dlk	B	116	334	6e-70			90.13	DELTA-CHYMOTRYPSIN; CHAIN: A, C; DELTA-CHYMOTRYPSIN; CHAIN: B, D;	HYDROLASE DELTA-CHYMOTRYPSIN, PEPTIDIC INHIBITOR, CHLOROMETHYL KETONE
1815	1dlk	B	140	376	1.2e-73			101.03	DELTA-CHYMOTRYPSIN; CHAIN: A, C; DELTA-CHYMOTRYPSIN; CHAIN: B, D;	HYDROLASE DELTA-CHYMOTRYPSIN, PEPTIDIC INHIBITOR, CHLOROMETHYL KETONE
1815	1doj	A	73	334	4.5e-63			80.24	ALPHA-THROMBIN; CHAIN: A; HIRUGEN; CHAIN: B; RWJ-51438; CHAIN: I;	HYDROLASE/HYDROLASE INHIBITOR THROMBIN, SERINE PROTEASE, ENZYME INHIBITION
1815	1doj	A	90	382	7.5e-71			92.95	ALPHA-THROMBIN; CHAIN: A; HIRUGEN; CHAIN: B; RWJ-51438; CHAIN: I;	HYDROLASE/HYDROLASE INHIBITOR THROMBIN, SERINE PROTEASE, ENZYME INHIBITION
1815	1ekb	B	112	334	3e-72			97.69	ENTEROPEPTIDASE; CHAIN: A; ENTEROPEPTIDASE; CHAIN: B; VAL-ASP-ASP-ASP-LYS PEPTIDE; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPsinogen ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR
1815	1ekb	B	129	369	1.5e-79			104.93	ENTEROPEPTIDASE;	HYDROLASE/HYDROLASE

1100

Table 5

SEQ ID NO:	PDB ID	CHAI N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL ID score	Compound	PDB annotation
									CHAIN: A; ENTEROPEPTIDASE; CHAIN: B; VAL-ASP-ASP-ASP-LYS PEPTIDE; CHAIN: C;	INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPSINOGEN ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR
1815	1elt		131	377	3e-73			84.92	ELASTASE; 1ELT 4 CHAIN: NULL, 1ELT 5	SERINE PROTEINASE
1815	1euf	A	120	334	1.5e-62			76.49	DUODENASE; CHAIN: A;	HYDROLASE BOVINE DUODENASE, SERINE PROTEASE, DUAL SPECIFICITY, 2 CRYSTAL STRUCTURE
1815	1f8	A	120	334	1.5e-59			88.45	NATURAL KILLER CELL PROTEASE 1; CHAIN: A; B; ECOTIN; CHAIN: C, E; ECOTIN; CHAIN: D, F;	HYDROLASE/HYDROLASE INHIBITOR GRANZYME B; COMPLEX (SERINE PROTEASE/INHIBITOR), PROTEASE SUBSTRATE 2 INTERACTIONS, BETA STRAND STRUCTURE, CHYMOTRYPSIN FOLD, 3 GRANZYME B, ECOTIN
1815	1f8	A	137	361	1.5e-60			94.52	NATURAL KILLER CELL PROTEASE 1; CHAIN: A; B; ECOTIN; CHAIN: C, E; ECOTIN; CHAIN: D, F;	HYDROLASE/HYDROLASE INHIBITOR GRANZYME B; COMPLEX (SERINE PROTEASE/INHIBITOR), PROTEASE SUBSTRATE 2 INTERACTIONS, BETA STRAND STRUCTURE, CHYMOTRYPSIN FOLD, 3 GRANZYME B, ECOTIN
1815	1fw	A	106	334	3e-68			77.63	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
1815	1fw	A	123	371	3e-74			93.93	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL

1101

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									ACROSIN LIGHT CHAIN; CHAIN: L	
1815	1fz	A	103	334	9e-69			78.10	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
1815	1fz	A	120	392	3e-75			94.78	BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
1815	1fni	A	132	334	7.5e-79			79.05	TRYPsin; CHAIN: A;	HYDROLASE SERINE PROTEASE, HYDROLASE
1815	1fni	A	140	365	3e-86			89.59	TRYPsin; CHAIN: A;	HYDROLASE SERINE PROTEASE, HYDROLASE
1815	1fxy	A	123	334	3e-69			77.08	COAGULATION FACTOR XA-TRYPsin CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE (PPACK) WITH CHAIN: L;	COMPLEX (PROTEASE/INHIBITOR) TRYPsin, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
1815	1gct	A	113	334	1.3e-69			90.30	HYDROLASE (SERINE PROTEINASE) GAMMA-*CHYMOTRYPsin *A (E.C.3.4.21.1) (\$P*H 7.0) IGCT 3	
1815	1gct	A	130	380	9e-74			99.87	HYDROLASE (SERINE PROTEINASE) GAMMA-*CHYMOTRYPsin *A (E.C.3.4.21.1) (\$P*H 7.0) IGCT 3	
1815	1hhb	A	25	192	7.5e-35			64.46	CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6),

1102

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
1815	1ikn	D	2	208	6e-40			55.70	NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
1815	1jbu	H	143	383	1.2e-65			85.70	COAGULATION FACTOR VII; CHAIN: H; COAGULATION FACTOR VII; CHAIN: I; PEPTIDE EXOSITE INHIBITOR A-183; CHAIN: X; MYOTROPHIN; CHAIN: NULL	HYDROLASE SERUM PROTHROMBIN CONVERSION ACCELERATOR; SERUM PROTHROMBIN CONVERSION ACCELERATOR; SHIFTED REGISTRATION, BETA-STRANDS
1815	1myo		23	142	3e-25			76.78	MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
1815	1ntf	E	1	200	3e-40			59.20	NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
1815	1pyt	D	104	334	3e-65			86.27	PROCARBOXYPEPTIDAS E A; CHAIN: A, B; PROPROTEINASE E; CHAIN: C; CHYMOTRYPSINOGEN C; CHAIN: D;	TERNARY COMPLEX (ZYMOGEN) TC, PCPA-TC; TERNARY COMPLEX (ZYMOGEN), SERINE PROTEINASE, C-TERMINAL 2 PEPTIDASE
1815	1pyt	D	116	379	3e-71			94.42	PROCARBOXYPEPTIDAS E A; CHAIN: A, B; PROPROTEINASE E; CHAIN: C; CHYMOTRYPSINOGEN C; CHAIN: D;	TERNARY COMPLEX (ZYMOGEN) TC, PCPA-TC; TERNARY COMPLEX (ZYMOGEN), SERINE PROTEINASE, C-TERMINAL 2 PEPTIDASE

1103

Table 5

SEQ ID NO:	PDB ID	CHAIN ID	STAR TAA	END AA	Psi Blast	Verify score	PMF score	SEQRESOL D score	Compound	PDB annotation
1815	1qnj	A	114	334	1.5e-62			79.44	ELASTASE; CHAIN: A;	HYDROLASE (SERINE PROTEASE) PPE; HYDROLASE(SERINE PROTEASE), ATOMIC RESOLUTION
1815	1qnj	A	131	380	1.4e-72			90.58	ELASTASE; CHAIN: A;	HYDROLASE (SERINE PROTEASE) PPE; HYDROLASE(SERINE PROTEASE), ATOMIC RESOLUTION
1815	1rft	A	28	148	3e-34	0.02-	-0.20		COAGULATION FACTOR IX; CHAIN: A;	COAGULATION FACTOR SERINE PROTEINASE, BLOOD COAGULATION, COAGULATION FACTOR
1815	1rft	B	111	334	6e-61			82.14	TWO CHAIN TISSUE PLASMINOGEN ACTIVATOR; CHAIN: A, B;	SERINE PROTEASE (TC)-T-PA; SERINE PROTEASE, FIBRINOLYTIC ENZYMES
1815	1rft	B	128	370	4.5e-72			89.37	TWO CHAIN TISSUE PLASMINOGEN ACTIVATOR; CHAIN: A, B;	SERINE PROTEASE (TC)-T-PA; SERINE PROTEASE, FIBRINOLYTIC ENZYMES
1815	1sgf	G	123	334	7.5e-71			79.09	NERVE GROWTH FACTOR; CHAIN: A, B, G, X, Y, Z;	GROWTH FACTOR 7S NGF; GROWTH FACTOR (BETA-NGF), HYDROLASE - SERINE PROTEINASE 2 (GAMMA-NGF), INACTIVE SERINE PROTEINASE (ALPHA-NGF)
1815	1sgf	G	140	362	6e-79			85.83	NERVE GROWTH FACTOR; CHAIN: A, B, G, X, Y, Z;	GROWTH FACTOR 7S NGF; GROWTH FACTOR (BETA-NGF), HYDROLASE - SERINE PROTEINASE 2 (GAMMA-NGF), INACTIVE SERINE PROTEINASE (ALPHA-NGF)
1815	1tm	A	123	334	4.5e-77			77.36	HYDROLASE (SERINE PROTEINASE) TRYPSIN	

Table 5

SEQ ID NO:	PDB ID	CHAIN N ID	STAR T AA	END AA	Psi Blast	Verify score	PMF score	SEQFOL D score	Compound	PDB annotation
									(E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR ITRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) ITRN 4 HUMAN TRYPSIN, DFP INHIBITED ITRN 6	
1815	1yes	B	25	229	7.5e-22			71.42	P53; CHAIN: A; 53BP2; CHAIN: B;	COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS) P53BP2; ANKYRIN REPEATS, SH3, P53, TUMOR SUPPRESSOR, MULTIGENE 2 FAMILY, NUCLEAR PROTEIN, PHOSPHORYLATION, DISEASE MUTATION, 3 POLYMORPHISM, COMPLEX (ANTI-ONCOGENE/ANKYRIN REPEATS)
1815	2sta	E	140	379	1.4e-85	0.04	-0.20		TRYPSIN; CHAIN: E; TRYPSIN INHIBITOR; CHAIN: I	HYDROLASE/HYDROLASE INHIBITOR SERINE PROTEINASE, TRYPSIN INHIBITOR
1817	1ds1	A	53	400	1.8e-46			52.55	CLAVAMINATE SYNTHASE I; CHAIN: A;	OXIDOREDUCTASE, LYASE CLAVAMINIC ACID SYNTHASE 1, CAS1; OXYGENASE, TRIFUNCTIONAL ENZYME, CLAVAMINATE SYNTHASE 1



Table 6  
1105

SEQ ID	Position	Maximum score	Average score
916	1-35	0.978	0.879
918	1-21	0.953	0.900
921	1-19	0.957	0.919
923	1-45	0.961	0.594
924	1-64	0.994	0.484
925	1-25	0.936	0.635
926	1-32	0.997	0.899
927	1-21	0.991	0.946
928	1-25	0.929	0.613
929	1-35	0.850	0.485
930	1-39	0.861	0.465
931	1-29	0.906	0.675
932	1-24	0.905	0.496
933	1-24	0.993	0.956
934	1-31	0.937	0.811
935	1-21	0.989	0.939
936	1-20	0.973	0.941
937	1-11	0.800	0.691
938	1-24	0.912	0.763
939	1-22	0.992	0.851
940	1-20	0.975	0.934
941	1-16	0.978	0.797
942	1-23	0.966	0.816
943	1-19	0.905	0.660
944	1-56	0.966	0.498
945	1-40	0.859	0.533
946	1-20	0.953	0.751
947	1-31	0.833	0.502
948	1-39	0.986	0.641
949	1-26	0.978	0.932
950	1-67	0.976	0.450
951	1-20	0.991	0.708
952	1-36	0.954	0.637
953	1-16	0.844	0.480
954	1-32	0.953	0.521
955	1-58	0.989	0.500
956	1-25	0.981	0.896
957	1-28	0.954	0.836
958	1-19	0.982	0.943
959	1-55	0.988	0.465
960	1-33	0.925	0.630
961	1-40	0.916	0.477
962	1-13	0.816	0.645
963	1-17	0.946	0.734
964	1-17	0.985	0.930
965	1-51	0.991	0.500
966	1-49	0.963	0.508
967	1-51	0.971	0.493
968	1-15	0.959	0.731
969	1-13	0.822	0.537
970	1-16	0.862	0.496
971	1-31	0.956	0.674
972	1-32	0.855	0.601
973	1-15	0.826	0.532
974	1-57	0.954	0.648

Table 6  
1106

SEQ ID	Position	Maximum score	Average score
975	1-20	0.976	0.854
976	1-22	0.873	0.586
977	1-26	0.985	0.839
978	1-49	0.963	0.774
979	1-26	0.843	0.609
980	1-57	0.966	0.561
981	1-16	0.912	0.705
982	1-22	0.962	0.827
983	1-13	0.803	0.580
984	1-51	0.932	0.462
985	1-20	0.982	0.890
986	1-20	0.977	0.923
987	1-21	0.977	0.868
988	1-23	0.838	0.580
989	1-39	0.905	0.628
990	1-23	0.989	0.864
991	1-20	0.884	0.650
992	1-19	0.985	0.920
993	1-28	0.887	0.505
994	1-27	0.955	0.743
995	1-27	0.992	0.934
996	1-53	0.953	0.723
997	1-30	0.905	0.522
998	1-20	0.954	0.746
999	1-19	0.821	0.482
1000	1-20	0.978	0.953
1001	1-45	0.890	0.608
1037	1-31	0.921	0.630
1038	1-36	0.972	0.563
1039	1-39	0.976	0.551
1040	1-26	0.937	0.703
1041	1-74	0.991	0.543
1042	1-23	0.945	0.797
1043	1-16	0.977	0.506
1044	1-21	0.967	0.759
1045	1-22	0.861	0.539
1046	1-27	0.934	0.682
1047	1-18	0.983	0.962
1048	1-22	0.827	0.517
1049	1-34	0.980	0.703
1050	1-15	0.987	0.955
1051	1-28	0.995	0.945
1052	1-18	0.995	0.977
1053	1-25	0.935	0.739
1054	1-19	0.976	0.950
1055	1-49	0.933	0.538
1056	1-22	0.918	0.723
1057	1-25	0.972	0.902
1058	1-52	0.981	0.622
1059	1-75	0.969	0.541
1060	1-75	0.979	0.817
1061	1-22	0.957	0.756
1062	1-45	0.978	0.852
1063	1-29	0.984	0.954
1064	1-36	0.994	0.713

Table 6  
1107

SEQ ID	Position	Maximum score	Average score
1065	1-34	0.914	0.608
1066	1-25	0.986	0.952
1067	1-38	0.990	0.909
1068	1-54	0.908	0.565
1069	1-20	0.870	0.708
1070	1-27	0.985	0.856
1071	1-23	0.966	0.812
1072	1-54	0.967	0.524
1160	1-36	0.948	0.776
1161	1-26	0.962	0.783
1162	1-31	0.919	0.543
1164	1-26	0.988	0.886
1166	1-53	0.998	0.751
1167	1-20	0.993	0.968
1168	1-24	0.987	0.924
1169	1-21	0.973	0.760
1170	1-56	0.962	0.585
1171	1-39	0.978	0.768
1172	1-25	0.969	0.937
1173	1-79	0.967	0.612
1174	1-23	0.990	0.934
1175	1-23	0.990	0.932
1176	1-42	0.977	0.677
1177	1-22	0.969	0.897
1178	1-17	0.869	0.641
1179	1-45	0.987	0.656
1180	1-31	0.956	0.853
1181	1-64	0.996	0.558
1182	1-50	0.986	0.566
1183	1-38	0.987	0.864
1184	1-28	0.987	0.924
1185	1-48	0.911	0.557
1186	1-58	0.914	0.462
1187	1-20	0.992	0.794
1188	1-24	0.894	0.608
1189	1-19	0.986	0.967
1190	1-23	0.965	0.817
1191	1-23	0.965	0.817
1192	1-53	0.884	0.504
1193	1-31	0.982	0.857
1194	1-18	0.805	0.583
1195	1-20	0.950	0.843
1196	1-25	0.803	0.575
1197	1-26	0.869	0.476
1198	1-23	0.982	0.866
1199	1-21	0.959	0.915
1200	1-30	0.983	0.652
1201	1-21	0.932	0.654
1202	1-21	0.932	0.654
1203	1-20	0.932	0.697
1204	1-26	0.946	0.775
1205	1-20	0.986	0.948
1206	1-64	0.865	0.475
1207	1-44	0.881	0.488
1208	1-19	0.991	0.973

Table 6  
1108

SEQ ID	Position	Maximum score	Average score
1209	1-38	0.938	0.459
1210	1-17	0.954	0.741
1211	1-20	0.916	0.802
1212	1-24	0.961	0.879
1213	1-20	0.938	0.716
1214	1-19	0.959	0.752
1215	1-55	0.946	0.682
1216	1-23	0.965	0.883
1217	1-23	0.965	0.883
1218	1-21	0.909	0.693
1219	1-17	0.971	0.886
1220	1-22	0.833	0.607
1221	1-16	0.948	0.836
1222	1-27	0.968	0.788
1223	1-23	0.853	0.523
1224	1-23	0.842	0.544
1225	1-19	0.937	0.713
1226	1-47	0.947	0.637
1227	1-21	0.984	0.933
1228	1-34	0.969	0.882
1229	1-15	0.894	0.745
1230	1-26	0.966	0.760
1231	1-28	0.883	0.474
1232	1-32	0.939	0.693
1233	1-28	0.915	0.655
1234	1-28	0.990	0.959
1235	1-39	0.926	0.696
1236	1-17	0.882	0.558
1237	1-16	0.967	0.933
1238	1-19	0.942	0.799
1239	1-19	0.910	0.773
1240	1-19	0.947	0.805
1241	1-20	0.976	0.875
1242	1-20	0.976	0.875
1243	1-15	0.973	0.942
1244	1-16	0.965	0.911
1245	1-17	0.985	0.956
1246	1-21	0.883	0.620
1247	1-13	0.874	0.527
1248	1-21	0.945	0.852
1249	1-24	0.903	0.540
1250	1-26	0.959	0.786
1251	1-30	0.934	0.585
1252	1-27	0.962	0.740
1253	1-25	0.842	0.507
1254	1-19	0.965	0.797
1255	1-20	0.970	0.804
1256	1-19	0.956	0.791
1257	1-21	0.977	0.868
1258	1-23	0.932	0.770
1259	1-31	0.960	0.635
1260	1-24	0.954	0.803
1261	1-43	0.949	0.467
1262	1-28	0.951	0.667
1263	1-28	0.963	0.844

Table 6  
1109

SEQ ID	Position	Maximum score	Average score
1264	1-60	0.971	0.472
1265	1-17	0.950	0.878
1266	1-22	0.984	0.852
1267	1-32	0.939	0.646
1268	1-50	0.939	0.491
1269	1-42	0.841	0.500
1270	1-23	0.923	0.602
1271	1-20	0.881	0.514
1272	1-20	0.881	0.514
1273	1-21	0.839	0.518
1274	1-33	0.955	0.547
1275	1-26	0.884	0.488
1276	1-27	0.867	0.555
1277	1-18	0.822	0.521
1278	1-20	0.814	0.562
1279	1-24	0.921	0.756
1280	1-50	0.845	0.490
1332	1-27	0.968	0.912
1335	1-38	0.937	0.565
1338	1-23	0.993	0.924
1339	1-28	0.995	0.945
1340	1-28	0.995	0.945
1345	1-73	0.997	0.744
1346	1-24	0.985	0.690
1347	1-16	0.983	0.952
1348	1-16	0.949	0.869
1349	1-16	0.996	0.959
1354	1-49	0.989	0.784
1355	1-77	0.963	0.578
1356	1-24	0.974	0.882
1357	1-38	0.979	0.810
1359	1-24	0.981	0.938
1361	1-24	0.998	0.972
1362	1-22	0.963	0.884
1364	1-33	0.997	0.864
1366	1-20	0.890	0.648
1367	1-23	0.996	0.932
1370	1-38	0.983	0.894
1371	1-19	0.960	0.521
1372	1-18	0.986	0.589
1373	1-20	0.929	0.768
1375	1-17	0.977	0.921
1377	1-19	0.985	0.935
1379	1-40	0.961	0.521
1382	1-21	0.985	0.922
1384	1-51	0.976	0.781
1385	1-29	0.984	0.954
1386	1-19	0.981	0.939
1387	1-17	0.979	0.938
1388	1-33	0.953	0.828
1390	1-42	0.974	0.616
1394	1-28	0.989	0.912
1396	1-33	0.979	0.698
1397	1-24	0.927	0.726
1399	1-46	0.991	0.727

Table 6  
1110

SEQ ID	Position	Maximum score	Average score
1400	1-44	0.900	0.657
1401	1-17	0.996	0.872
1402	1-39	0.989	0.941
1403	1-21	0.975	0.604
1404	1-16	0.957	0.870
1405	1-20	0.990	0.875
1407	1-71	0.981	0.566
1408	1-18	0.985	0.928
1410	1-14	0.941	0.861
1412	1-32	0.935	0.549
1419	1-20	0.989	0.960
1420	1-50	0.987	0.583
1421	1-27	0.977	0.849
1422	1-45	0.981	0.817
1423	1-28	0.988	0.923
1430	1-73	0.997	0.731
1432	1-23	0.992	0.956
1433	1-23	0.992	0.956
1437	1-60	0.976	0.737
1442	1-27	0.975	0.937
1467	1-57	0.931	0.514
1468	1-15	0.810	0.544
1469	1-41	0.926	0.451
1470	1-31	0.914	0.618
1471	1-33	0.889	0.542
1472	1-20	0.941	0.754
1473	1-55	0.929	0.541
1474	1-36	0.951	0.625
1475	1-18	0.897	0.608
1476	1-22	0.980	0.917
1477	1-73	0.959	0.450
1478	1-18	0.946	0.791
1479	1-22	0.951	0.739
1480	1-73	0.968	0.456
1481	1-25	0.879	0.674
1482	1-20	0.946	0.806
1483	1-24	0.992	0.825
1484	1-17	0.836	0.620
1485	1-53	0.917	0.508
1486	1-19	0.872	0.724
1487	1-58	0.954	0.452
1488	1-19	0.966	0.927
1489	1-36	0.979	0.862
1490	1-25	0.976	0.803
1491	1-26	0.898	0.658
1492	1-22	0.948	0.844
1493	1-22	0.894	0.673
1494	1-52	0.937	0.660
1495	1-37	0.930	0.614
1496	1-25	0.979	0.914
1497	1-22	0.972	0.929
1498	1-21	0.978	0.816
1499	1-24	0.839	0.692
1500	1-15	0.823	0.560
1501	1-16	0.960	0.845

Table 6  
1111

SEQ ID	Position	Maximum score	Average score
1502	1-29	0.940	0.683
1503	1-12	0.832	0.523
1504	1-36	0.945	0.749
1505	1-33	0.843	0.488
1506	1-23	0.901	0.626
1507	1-20	0.834	0.505
1508	1-25	0.919	0.690
1509	1-44	0.885	0.513
1510	1-23	0.974	0.825
1511	1-13	0.941	0.847
1512	1-42	0.885	0.513
1513	1-18	0.903	0.592
1514	1-30	0.948	0.816
1515	1-16	0.944	0.818
1516	1-23	0.849	0.631
1517	1-13	0.813	0.610
1518	1-20	0.962	0.921
1519	1-23	0.823	0.495
1520	1-2	0.893	0.458
1521	1-17	0.902	0.762
1522	1-20	0.952	0.772
1523	1-23	0.991	0.925
1524	1-15	0.858	0.647
1525	1-15	0.979	0.961
1526	1-23	0.926	0.669
1527	1-23	0.942	0.686
1528	1-22	0.967	0.923
1529	1-22	0.913	0.519
1530	1-27	0.867	0.508
1531	1-15	0.807	0.631
1532	1-39	0.907	0.475
1533	1-24	0.823	0.503
1534	1-16	0.892	0.668
1535	1-16	0.860	0.607
1536	1-28	0.920	0.686
1537	1-20	0.801	0.618
1538	1-28	0.887	0.643
1539	1-21	0.927	0.568
1540	1-22	0.997	0.951
1541	1-22	0.985	0.827
1542	1-24	0.962	0.871
1543	1-22	0.972	0.932
1544	1-23	0.908	0.678
1545	1-27	0.899	0.579
1546	1-16	0.988	0.955
1547	1-16	0.873	0.622
1548	1-28	0.840	0.493
1549	1-15	0.889	0.689
1550	1-21	0.889	0.633
1551	1-44	0.981	0.617
1552	1-18	0.881	0.649
1553	1-16	0.805	0.631
1554	1-22	0.965	0.807
1555	1-26	0.928	0.744
1556	1-37	0.934	0.475

Table 6  
1112

SEQ ID	Position	Maximum score	Average score
1557	1-31	0.935	0.631
1558	1-20	0.990	0.868
1559	1-13	0.853	0.530
1560	1-21	0.833	0.604
1561	1-25	0.922	0.642
1562	1-24	0.960	0.683
1563	1-31	0.908	0.677
1564	1-22	0.980	0.937
1565	1-17	0.857	0.523
1566	1-15	0.878	0.521
1567	1-19	0.861	0.660
1568	1-18	0.963	0.812
1569	1-32	0.965	0.659
1570	1-21	0.919	0.457
1571	1-33	0.931	0.799
1572	1-11	0.962	0.656
1573	1-22	0.969	0.862
1574	1-18	0.907	0.508
1575	1-19	0.952	0.751
1576	1-16	0.949	0.807
1577	1-20	0.984	0.912
1578	1-47	0.956	0.583
1579	1-30	0.965	0.639
1580	1-33	0.921	0.650
1581	1-21	0.996	0.915
1582	1-18	0.992	0.969
1583	1-19	0.820	0.492
1584	1-33	0.895	0.469
1585	1-37	0.987	0.701
1586	1-19	0.882	0.606
1587	1-19	0.896	0.652
1588	1-11	0.896	0.803
1589	1-22	0.977	0.769
1590	1-21	0.916	0.742
1591	1-15	0.909	0.760
1592	1-13	0.925	0.649
1593	1-26	0.963	0.782
1594	1-13	0.840	0.530
1595	1-35	0.984	0.835
1596	1-17	0.813	0.642
1597	1-20	0.822	0.535
1598	1-22	0.996	0.965
1599	1-17	0.866	0.603
1600	1-28	0.983	0.873
1601	1-36	0.963	0.813
1602	1-16	0.942	0.750
1603	1-16	0.827	0.545
1604	1-19	0.990	0.935
1605	1-28	0.910	0.588
1606	1-28	0.990	0.772
1607	1-30	0.944	0.835
1608	1-36	0.963	0.842
1609	1-18	0.937	0.839
1610	1-25	0.966	0.910
1611	1-33	0.963	0.577



Table 6  
1113

SEQ ID	Position	Maximum score	Average score
1612	1-27	0.985	0.755
1613	1-20	0.970	0.737
1614	1-11	0.872	0.706
1615	1-19	0.932	0.541
1616	1-14	0.987	0.968
1617	1-52	0.948	0.570
1618	1-31	0.842	0.473
1619	1-27	0.985	0.927
1620	1-17	0.822	0.503
1621	1-61	0.982	0.770
1622	1-20	0.876	0.549
1623	1-19	0.872	0.552
1624	1-16	0.928	0.878
1625	1-20	0.914	0.495
1626	1-17	0.874	0.653
1627	1-16	0.969	0.917
1628	1-21	0.886	0.568
1629	1-45	0.931	0.487
1630	1-25	0.937	0.771
1631	1-23	0.811	0.514
1632	1-21	0.923	0.724
1633	1-17	0.833	0.543
1634	1-24	0.979	0.927
1635	1-20	0.972	0.802
1636	1-22	0.932	0.577
1637	1-18	0.894	0.643
1638	1-29	0.952	0.789
1639	1-38	0.918	0.617
1640	1-26	0.848	0.707
1641	1-23	0.989	0.864
1642	1-67	0.995	0.642
1643	1-18	0.970	0.681
1644	1-16	0.982	0.941
1645	1-24	0.971	0.908
1646	1-25	0.969	0.794
1647	1-40	0.863	0.548
1648	1-20	0.984	0.822
1649	1-15	0.803	0.568
1650	1-57	0.986	0.558
1651	1-26	0.975	0.763
1652	1-19	0.996	0.953
1653	1-34	0.994	0.961
1654	1-37	0.973	0.869
1655	1-21	0.901	0.716
1656	1-21	0.990	0.932
1657	1-42	0.959	0.693
1658	1-18	0.928	0.723
1659	1-16	0.914	0.756
1660	1-16	0.841	0.615
1661	1-70	0.932	0.496
1662	1-25	0.835	0.567
1663	1-39	0.958	0.854
1664	1-19	0.977	0.790
1665	1-19	0.962	0.794
1666	1-22	0.991	0.954

Table 6  
1114

SEQ ID	Position	Maximum score	Average score
1667	1-39	0.842	0.525
1668	1-24	0.907	0.722
1669	1-52	0.950	0.667
1670	1-17	0.940	0.583
1671	1-26	0.926	0.642
1672	1-20	0.964	0.902
1673	1-22	0.947	0.788
1674	1-27	0.889	0.642
1675	1-16	0.985	0.900
1676	1-42	0.869	0.503
1677	1-73	0.976	0.512
1678	1-21	0.951	0.816
1679	1-25	0.946	0.771
1680	1-18	0.975	0.954
1681	1-19	0.926	0.729
1682	1-16	0.975	0.891
1683	1-25	0.967	0.881
1684	1-30	0.877	0.492
1685	1-23	0.990	0.905
1686	1-16	0.985	0.964
1687	1-20	0.969	0.803
1688	1-23	0.965	0.829
1689	1-20	0.894	0.502
1690	1-23	0.884	0.582
1691	1-19	0.972	0.705
1692	1-16	0.887	0.670
1693	1-21	0.961	0.741
1694	1-16	0.913	0.452
1695	1-25	0.946	0.829
1696	1-44	0.985	0.510
1697	1-32	0.920	0.468
1698	1-38	0.886	0.519
1699	1-15	0.978	0.904
1700	1-20	0.930	0.666
1701	1-25	0.988	0.903
1702	1-26	0.901	0.507
1703	1-38	0.971	0.793
1704	1-20	0.977	0.888
1705	1-30	0.822	0.550
1706	1-22	0.844	0.472
1707	1-16	0.923	0.641
1708	1-21	0.906	0.624
1709	1-22	0.977	0.880
1710	1-19	0.946	0.798
1711	1-24	0.928	0.557
1712	1-27	0.969	0.901
1713	1-30	0.926	0.755
1714	1-20	0.975	0.913
1715	1-23	0.927	0.606
1716	1-18	0.924	0.715
1717	1-31	0.826	0.618
1718	1-33	0.977	0.811
1719	1-26	0.934	0.636
1720	1-38	0.960	0.483
1721	1-36	0.962	0.537

Table 6  
1115

SEQ ID	Position	Maximum score	Average score
1722	1-70	0.993	0.825
1723	1-33	0.928	0.642
1724	1-38	0.967	0.807
1725	1-18	0.988	0.897
1726	1-16	0.974	0.949
1727	1-42	0.951	0.468
1728	1-29	0.904	0.595
1729	1-41	0.938	0.600
1730	1-22	0.829	0.545
1731	1-52	0.990	0.466
1732	1-26	0.970	0.909
1733	1-24	0.971	0.865
1734	1-39	0.996	0.867
1735	1-45	0.961	0.593
1736	1-22	0.991	0.950
1737	1-25	0.978	0.623
1738	1-18	0.991	0.947
1739	1-47	0.926	0.486
1740	1-20	0.985	0.959
1741	1-32	0.995	0.972
1742	1-17	0.958	0.656
1743	1-62	0.964	0.638
1744	1-18	0.892	0.717
1745	1-27	0.992	0.934
1746	1-21	0.989	0.962
1747	1-67	0.991	0.685
1748	1-22	0.898	0.623
1749	1-15	0.981	0.945
1750	1-19	0.971	0.786
1751	1-36	0.983	0.664
1752	1-30	0.988	0.830
1753	1-28	0.953	0.866
1754	1-64	0.967	0.707
1755	1-56	0.967	0.554
1756	1-35	0.945	0.589
1757	1-19	0.922	0.737
1758	1-30	0.990	0.739
1759	1-25	0.951	0.864
1760	1-55	0.990	0.483
1761	1-29	0.982	0.792
1762	1-27	0.936	0.759
1763	1-29	0.960	0.805
1764	1-18	0.978	0.906
1765	1-22	0.962	0.920
1766	1-21	0.997	0.955
1767	1-25	0.942	0.847
1768	1-32	0.989	0.830
1769	1-25	0.880	0.630
1770	1-26	0.872	0.514
1771	1-29	0.990	0.648
1772	1-38	0.901	0.670
1773	1-25	0.992	0.955
1774	1-22	0.839	0.574
1775	1-33	0.970	0.887
1776	1-43	0.987	0.716

Table 6  
1116

SEQ ID	Position	Maximum score	Average score
1777	1-29	0.831	0.548
1778	1-42	0.992	0.654
1779	1-20	0.936	0.580
1780	1-24	0.985	0.941
1781	1-48	0.996	0.882
1782	1-58	0.973	0.595
1783	1-22	0.899	0.602
1784	1-46	0.958	0.672
1785	1-10	0.881	0.707
1786	1-22	0.818	0.577
1787	1-53	0.903	0.465
1788	1-77	0.976	0.483
1789	1-72	0.973	0.475
1790	1-79	0.972	0.734
1791	1-71	0.958	0.519
1792	1-25	0.996	0.966
1793	1-75	0.975	0.571
1794	1-43	0.853	0.587
1795	1-17	0.948	0.513
1796	1-26	0.981	0.938
1797	1-33	0.986	0.753
1798	1-21	0.866	0.619
1799	1-31	0.965	0.885
1800	1-51	0.989	0.477
1801	1-65	0.977	0.706
1802	1-21	0.994	0.891
1803	1-24	0.986	0.956
1804	1-25	0.880	0.602
1805	1-24	0.991	0.954
1806	1-24	0.914	0.580
1807	1-71	0.956	0.528
1808	1-38	0.901	0.670
1809	1-43	0.975	0.765
1810	1-14	0.814	0.590
1811	1-14	0.909	0.490
1812	1-54	0.990	0.503
1813	1-76	0.999	0.992
1814	1-61	0.912	0.473
1815	1-20	0.975	0.944
1816	1-45	0.983	0.465
1817	1-24	0.936	0.769
1818	1-62	0.997	0.748
1819	1-74	0.986	0.539
1820	1-59	0.959	0.559
1821	1-22	0.860	0.539
1822	1-30	0.996	0.822

Table 7  
1117

SEQ ID	Chromosomal location
1	13q12-q14
2	9q33-q34
3	2
4	13q12-q14
5	13q12-q14
6	13q12-q14
7	13q12-q14
8	13q12-q14
9	13q12-q14
10	17q11.2
11	19q13.3-q13.4
12	2
13	13q12-q14
14	8q
15	8q
16	11q13.1-q13.3
17	20q12
18	20q12
19	20q12
20	4
21	4
22	9q33-q34
23	9q33-q34
24	9q33-q34
25	9q33-q34
26	9q33-q34
27	9q33-q34
28	9q33-q34
29	9q33-q34
30	9q33-q34
31	9q33-q34
32	2
33	2
34	13q12-q14
35	13q12-q14
36	13q12-q14
37	13q12-q14
38	13q12-q14
39	13q12-q14
40	13q12-q14
41	13q12-q14
42	13q12-q14
43	13q12-q14
44	13q12-q14
45	13q12-q14
46	13q12-q14
47	13q12-q14
48	13q12-q14
49	13q12-q14
50	13q12-q14
51	13q12-q14
52	13q12-q14
53	13q12-q14
54	13q12-q14
55	13q12-q14

Table 7  
1118

SEQ ID	Chromosomal location
56	13q12-q14
57	13q12-q14
58	13q12-q14
59	13q12-q14
60	13q12-q14
61	13q12-q14
62	13q12-q14
63	13q12-q14
64	13q12-q14
65	13q12-q14
66	13q12-q14
67	13q12-q14
68	13q12-q14
69	13q12-q14
70	13q12-q14
71	13q12-q14
72	13q12-q14
73	13q12-q14
74	13q12-q14
75	13q12-q14
76	13q12-q14
77	2
78	7
79	2
80	2
81	19q13.3-q13.4
82	19q13.3-q13.4
83	19q13.3-q13.4
84	19q13.3-q13.4
85	19q13.3-q13.4
86	19q13.3-q13.4
87	19q13.3-q13.4
88	19q13.3-q13.4
89	19q13.3-q13.4
90	19q13.3-q13.4
91	9q33-q34
92	13q12-q14
93	20q12
94	7
95	4
96	4
97	4
98	9q33-q34
99	9q33-q34
100	9q33-q34
101	9q33-q34
102	9q33-q34
103	9q33-q34
104	9q33-q34
105	2
106	13q12-q14
107	13q12-q14
108	13q12-q14
109	13q12-q14
110	13q12-q14

Table 7  
1119

SEQ ID	Chromosomal location
111	13q12-q14
112	13q12-q14
113	13q12-q14
114	13q12-q14
115	13q12-q14
116	13q12-q14
117	13q12-q14
118	13q12-q14
119	13q12-q14
120	2
121	2
122	2
123	19q13.3-q13.4
124	19q13.3-q13.4
125	19q13.3-q13.4
126	13q12-q14
127	13q12-q14
128	20q12
129	4
130	4
131	9q33-q34
132	2
133	2
134	2
135	13q12-q14
136	13q12-q14
137	13q12-q14
138	13q12-q14
139	13q12-q14
140	17
141	6p21.3
142	13q12-q14
143	13q12-q14
144	13q12-q14
145	13q12-q14
146	13q12-q14
147	13q12-q14
148	13q12-q14
149	13q12-q14
150	13q12-q14
151	13q12-q14
152	13q12-q14
153	13q12-q14
154	13q12-q14
155	13q12-q14
156	2
157	2
158	19q13.3-q13.4
159	19q13.3-q13.4
160	19q13.3-q13.4
161	19q13.3-q13.4
162	4
163	4
164	9q33-q34
165	13q12-q14

Table 7  
1120

SEQ ID	Chromosomal location
166	13q12-q14
167	13q12-q14
168	13q12-q14
169	22q12.1-12.3
170	13q12-q14
171	13q12-q14
172	13q12-q14
173	13q12-q14
174	13q12-q14
175	13q12-q14
176	8q
177	20q12
178	20q12
179	4
180	4
181	4
182	9q33-q34
183	9q33-q34
184	9q33-q34
185	3
186	9q33-q34
187	9q33-q34
188	9q33-q34
189	9q33-q34
190	9q33-q34
191	9q33-q34
192	9q33-q34
193	2
194	2
195	13q12-q14
196	13q12-q14
197	13q12-q14
198	13q12-q14
199	13q12-q14
200	13q12-q14
201	13q12-q14
202	13q12-q14
203	13q12-q14
204	13q12-q14
205	13q12-q14
206	13q12-q14
207	13q12-q14
208	13q12-q14
209	13q12-q14
210	13q12-q14
211	13q12-q14
212	13q12-q14
213	13q12-q14
214	19q13
215	13q12-q14
216	13q12-q14
217	13q12-q14
218	13q12-q14
219	13q12-q14
220	13q12-q14



Table 7  
1121

SEQ ID	Chromosomal location
221	13q12-q14
222	13q12-q14
223	13q12-q14
224	13q12-q14
225	13q12-q14
226	13q12-q14
227	13q12-q14
228	13q12-q14
229	13q12-q14
230	13q12-q14
231	13q12-q14
232	13q12-q14
233	13q12-q14
234	13q12-q14
235	13q12-q14
236	13q12-q14
237	13q12-q14
238	5
239	5
240	19q13.3-q13.4
241	19q13.3-q13.4
242	19q13.3-q13.4
243	19q13.3-q13.4
244	19q13.3-q13.4
245	19q13.3-q13.4
246	19q13.3-q13.4
247	19q13.3-q13.4
248	19q13.3-q13.4
250	6
256	1p32
259	11q12-q13.1
268	18p11.22-p11.21
269	7q11.23
270	3
272	9q34.3
278	9q32-q33
279	14q11.2
280	14q11.2
283	2
285	13
298	20q12-q13.1
303	1
305	19
306	19
310	4p16-p15
311	6
316	3p22-p21.3
326	1p31.1-p22.3
335	11pter-p15.5
336	3
350	11q
353	3
355	8p21-p12
356	1q23-q25.1
357	1q23-q25.1

Table 7  
1122

SEQ ID	Chromosomal location
359	20q13.2-q13.33
360	13
361	13
362	22q13.1
369	6
373	2p24.3-p24.1
374	2p24.3-p24.1
383	11q13
384	19q13.2
388	17q11-q21.1
398	8p23.2-p23.1
403	19
407	11
421	19q13.1
427	20q12-q13.12
428	17
429	17
431	11q13
432	7q35
433	X
435	6q22.2-22.33
437	19
439	16q24.3
445	19q13
446	12
452	14
453	3
456	14
457	10
458	10
459	3p21.1-q13.13
460	1p36
470	14q31
475	4q11-q13
476	6
477	15q21-q22
478	15q24-q26
493	19
496	6q24.1-24.3
497	15
501	11q13
505	1q21
506	6p21.3
507	3
510	1q21.3
512	14
513	9
515	1
521	7q32-34
522	7q32-34
532	2
545	q21.2-21.3
547	11
551	11q22
556	10,

Table 7  
1123

SEQ ID	Chromosomal location
557	9p24.1-24.3,
559	16
560	18
561	9
562	22q11.22-12.3.
565	6q22.3-23.3
566	6q22.1-22.33,
567	1
568	11
569	6q21
571	18,
572	15,
573	5
574	5
575	6p21.1-21.2.
576	8,
577	10
578	1
579	12
580	1
581	6q16.3-22.1,
582	3
583	10
585	3
587	1
588	5
590	4
592	1q24.
593	1
594	8
595	17,
596	11
597	5
598	16
599	11
600	13,
601	11q
603	2
604	6q25.2-26,
606	1
607	15
608	9
609	1q24
610	1
611	4
612	17,
613	8
614	17,
615	15
616	6,
617	15
618	2
619	3
620	17
621	8

Table 7  
1124

SEQ ID	Chromosomal location
622	14
623	6q24.1-25.2,
624	2
625	9
626	8
627	12q24
629	20,
630	NA
631	2
632	2
633	Xp11.21-11.22
634	17
636	2
637	22
638	1
639	3
640	19,
641	1
642	7
643	5
645	17
647	1,
649	6
650	7
651	5
652	19,
653	20q11.1-11.23,
654	1
655	3
656	8
657	16
659	9
660	18
662	6p21.3,
665	X
666	5
667	1
669	2
670	13,
671	1
672	15
673	9,
674	10
675	9,
676	7
678	16
679	X
680	11,
681	19,
683	8
684	4
685	6q16.1-16.3,
686	10,
687	14
688	4

Table 7  
1125

SEQ ID	Chromosomal location
689	6q26-27,
690	1p35.1-36.13,
692	6q26-27,
695	12
700	9,
701	5
702	2
703	8
704	10,
705	19
706	14
708	19,
709	17
711	16
712	16
713	5
714	16
715	19
716	16
717	12
718	15
719	15
720	14
721	16
722	5
723	6,
724	12q
725	1
726	8
727	15,
728	2
729	7
730	17,
731	21q22.1,
732	20p12
733	11
734	1
735	15
736	2
737	1p36.21-36.33.
738	1
739	6q21-22.31,
740	15
742	12
743	UL
744	13q33.3-34,
745	17,
746	12
747	2
749	16
752	17
754	16
755	20q12-13.12.
756	19
757	14

Table 7  
1126

SEQ ID	Chromosomal location
758	8,
759	6,
760	22
761	19
763	13,
764	1q32.1-41
765	8,
767	20
769	11q,
770	1
771	6p22.3-24.1,
772	9p23-24.3,
773	17
774	19
775	3
776	3
777	22q12
779	1p36.13-36.31,
780	17
781	7
783	4
784	6,
785	1
786	21q22.1,
787	6
788	1
789	15
790	20,
791	9
793	10,
794	16
796	18
797	18
798	10
799	8,
800	16
801	UNK
803	9,
807	20q13.1.
808	10,
809	3
810	UNK
815	20.
818	1q25.1-31.1,
820	10
821	20p12.2-13,
822	20
823	4,
824	1
826	5
827	8
829	7
830	1p21.3-22.3,
832	1
833	12q

Table 7  
1127

SEQ ID	Chromosomal location
834	6
836	6,
837	10
838	9p11.2-21,
839	11
840	5
841	17
842	11q
844	1
845	11q,
846	9
848	13,
849	12
852	1
853	16
854	15
856	6q25.3-26
857	6,
858	1
859	11
860	5p,
861	11
862	6,
864	1q22-q23,
865	12p
866	1q32.2-41
867	2
868	13q33.1-34,
869	6p21.3,
870	3
872	1
873	3
874	19,
876	9,
877	16p13
878	5
879	16
880	5
881	16
882	15
883	15
884	9
885	8p11.2,
886	17
887	5
888	6,
889	13,
890	11
891	11
892	Xq22.3-23,
894	9,
895	10
896	15
897	11
898	1,

Table 7  
1128

SEQ ID	Chromosomal location
899	6q25.2-26,
900	12
901	1
902	19
903	19,
904	4
905	5
906	5
908	19
909	4,
910	19,
911	15



Table 8  
1129

SEQ ID NO:	Number of Transmembrane Domains	For each Transmembrane Domain, its Transmembrane Domain Position in SEQ ID NO: and its TM Pred Score
1037	1	100-115:1952
1038	1	732-749:2593
1039	1	181-201:2410
1040	2	53-68:1828 132-149:2533
1041	2	53-69:2959 121-140:2878
1042	1	536-560:2906
1043	1	63-82:2545
1044	5	86-102:1766 189-205:2721 229-244:1878 273-300:1714 385-405:1946
1045	1	629-645:2430
1046	1	59-75:2149
1047	1	306-332:2773
1048	1	118-136:2329
1049	2	98-113:2861 220-243:2391
1050	1	151-169:2618
1051	2	94-110:2524 124-146:2138
1052	1	73-87:2180
1053	1	206-226:2584
1054	1	402-419:2096
1055	1	343-361:1953
1056	1	132-154:2199
1057	1	590-613:2402
1058	2	89-105:1748 155-173:2433
1059	1	201-222:2190
1060	4	254-277:2256 317-332:1771 442-460:2005 530-544:2110
1061	2	169-186:1866 239-259:2042
1062	1	63-77:1794
1063	1	227-248:3456
1064	1	133-148:2558
1065	3	435-453:1849 505-526:2495 697-712:2057
1066	1	317-340:2214
1067	1	173-192:2637
1068	1	63-79:1933
1069	1	124-146:3384
1070	3	82-102:2213 115-135:1769 160-185:2317
1071	1	312-329:2354
1072	2	116-131:3056 188-209:2254
1073	6	48-71:1708 174-196:2300 237-254:1918 359-378:1887 413-435:1864 501-518:2625
1074	6	136-159:1708 262-284:2300 325-342:1918 447-466:1887 501-523:1864 589-606:2625
1075	1	352-376:2946
1076	3	106-125:2854 226-241:1973 277-300:2759
1077	4	85-105:2047 208-225:1907 309-330:2122 454-471:2461
1079	1	60-75:2189
1081	3	192-214:1705 236-259:1933 436-453:2349
1082	1	459-477:1896
1083	1	144-159:3028
1084	1	144-159:3028
1085	1	436-455:2525
1086	1	705-724:2525
1093	2	93-108:2014 249-264:2324
1094	5	145-165:2633 316-331:2180 399-412:1770 481-496:2328 541-560:2589

Table 8  
1130

SEQ ID NO:	Number of Transmembrane Domains	For each Transmembrane Domain, its Transmembrane Domain Position in SEQ ID NO: and its TM Pred Score
1095	1	73-92:1951
1097	1	85-103:2195
1099	1	160-176:2085
1100	1	256-272:2085
1101	1	210-226:2085
1102	1	63-94:3259
1103	2	184-201:2183 245-262:1812
1104	1	276-295:3080
1106	1	179-194:2620
1110	1	111-129:2519
1112	1	169-190:2680
1113	3	61-82:2141 99-134:1715 119-139:2765
1114	1	65-85:1713
1115	3	123-137:2644 190-218:2074 300-314:2588
1116	2	98-123:2354 270-295:2148
1117	1	77-92:1791
1118	1	68-88:2672
1119	1	1524-1547:2939
1121	1	95-113:2958
1123	2	92-107:1923 162-178:2760
1124	1	71-94:1835
1125	1	379-403:3221
1127	1	152-182:1795
1129	3	201-217:2437 338-353:1761 449-466:2589
1130	2	99-114:1754 108-130:2731
1131	1	1177-1193:3038
1132	1	95-111:2301
1133	2	205-227:1715 307-322:1735
1134	1	308-330:2431
1136	5	92-107:1734 298-311:2063 363-378:1720 382-399:1988 453-471:2040
1137	2	56-75:2481 127-148:2269
1138	1	228-251:1822
1139	4	97-115:1903 177-194:1919 889-905:2063 988-1008:2027
1141	1	223-242:2971
1143	3	63-86:2169 177-194:1878 271-288:2186
1144	2	116-136:2390 179-194:2530
1146	2	66-82:2701 110-126:1755
1147	2	74-106:2580 139-156:1958
1148	1	522-544:2644
1149	2	83-97:2024 200-216:2275
1150	1	200-216:2275
1151	1	92-109:2588
1152	2	145-169:1834 317-346:1891
1155	4	64-83:2948 218-232:2016 452-480:1829 535-553:1999
1156	1	311-330:2524
1157	1	78-111:2597
1159	1	163-180:2270
1330	1	264-282:2674
1331	1	105-128:2620
1332	1	141-164:3091
1333	1	394-409:1914
1334	1	390-410:1947
1335	1	67-82:1917

Table 8  
1131

SEQ ID NO:	Number of Transmembrane Domains	For each Transmembrane Domain, its Transmembrane Domain Position in SEQ ID NO: and its TM Pred Score
1336	1	301-317:3356
1337	1	301-317:3356
1338	1	427-443:1894
1339	2	103-119:2524 133-155:2138
1340	2	94-110:2524 124-146:2138
1341	2	61-78:1768 212-233:2126
1342	1	139-160:2023
1343	1	331-353:1902
1344	1	99-117:2054
1345	1	52-67:1997
1346	1	67-80:2411
1347	1	367-384:2888
1348	1	457-476:2898
1349	2	202-217:1923 272-288:2760
1350	5	266-284:2563 328-346:2525 390-408:2516 452-470:2456 514-532:2395
1351	1	94-116:2763
1352	2	45-62:2164 71-85:1838
1353	1	482-504:3155
1354	3	1-17:2284 75-90:1824 145-160:2079
1355	1	12-28:1796
1356	1	332-356:3221
1357	1	9-30:2498
1358	3	202-215:2016 256-281:1796 283-312:1900
1359	1	267-284:3050
1360	3	229-245:2437 316-333:1885 429-446:2589
1361	1	1-17:2914
1362	1	4-21:2014
1363	1	962-980:1762
1364	1	12-31:2460
1365	1	38-54:3201
1366	1	103-119:2300
1367	1	1-15:2225
1368	1	48-64:1711
1369	1	48-64:1711
1370	1	21-36:2592
1371	2	290-307:2420 419-433:2394
1372	1	66-81:2345
1373	1	3-18:1914
1374	1	110-124:2032
1375	1	3-18:1862
1376	4	37-56:2540 102-127:1708 194-209:2784 434-453:2036
1377	1	1-20:2264
1378	3	259-274:2102 330-358:2448 416-434:2702
1379	1	24-43:2348
1380	1	196-211:1704
1381	3	159-176:1865 238-260:2690 305-319:2565
1382	1	262-288:2789
1383	2	60-89:2283 119-140:1980
1384	1	21-41:2833
1385	1	227-248:3456
1386	1	3-19:2081
1387	1	162-180:3181
1388	1	12-27:2098

Table 8  
1132

SEQ ID NO:	Number of Transmembrane Domains	For each Transmembrane Domain, its Transmembrane Domain Position in SEQ ID NO: and its TM Pred Score
1389	1	487-503:2039
1390	2	95-111:2234 155-173:2180
1391	2	87-106:2547 186-206:1903
1392	3	142-162:2149 169-184:1733 265-283:2036
1393	3	83-110:1977 142-159:1729 238-254:2361
1394	1	12-30:2458
1395	1	61-76:2746
1396	1	14-31:1963
1397	1	1060-1079:2742
1398	1	220-244:2016
1399	1	17-36:2664
1400	1	21-41:1851
1401	1	37-69:2282
1402	1	24-60:1926
1403	1	55-75:2438
1404	1	237-252:2493
1405	1	1-20:1866
1406	1	172-190:1781
1407	1	90-111:2881
1408	1	6-23:2294
1409	1	477-499:2659
1410	1	231-249:2797
1411	3	64-84:1993 95-110:1932 149-167:2173
1412	3	166-182:1937 216-239:1819 326-341:1779
1413	1	154-169:2224
1414	1	241-254:1760
1415	2	71-97:1818 234-254:2475
1416	1	143-164:2551
1417	2	104-122:1815 147-162:2031
1418	2	121-135:2273 209-234:2186
1419	2	8-27:2032 82-102:1765
1420	1	26-47:2557
1421	3	18-34:1798 220-243:1742 301-318:1915
1422	1	27-42:3053
1423	1	6-21:1974
1424	1	345-360:2050
1425	1	44-65:2195
1426	1	763-779:1916
1427	1	448-469:2686
1428	1	597-614:2205
1429	1	194-210:2157
1430	2	22-44:2983 52-67:1999
1431	2	22-39:1830 91-106:2145
1432	1	6-26:2422
1433	1	6-26:2422
1434	1	72-89:1942
1435	2	128-146:1770 337-352:3007
1436	1	88-105:1803
1437	1	1013-1035:2346
1438	4	574-589:1938 650-668:2275 720-735:1944 793-810:1729
1439	4	655-670:1938 731-749:2275 801-816:1944 874-891:1729
1440	1	65-81:2339
1441	1	131-151:1772
1442	2	5-22:1872 248-273:2467

Table 8  
1133

SEQ ID NO:	Number of Transmembrane Domains	For each Transmembrane Domain, its Transmembrane Domain Position in SEQ ID NO: and its TM Pred Score
1443	3	179-194:2318 387-403:3014 442-465:1871
1444	2	58-74:2684 238-264:2576
1720	1	89-107:1832
1721	1	225-240:2165
1722	2	13-29:2422 49-63:2158
1723	1	129-144:1704
1724	1	12-28:1900
1725	4	1-22:1840 143-158:1938 181-198:2860 247-265:2284
1726	1	1-15:1790
1727	2	100-122:2447 199-217:2147
1728	1	12-28:1802
1729	1	21-37:1706
1730	1	328-343:2384
1731	3	33-60:2908 185-211:2596 237-258:1818
1732	1	279-309:1795
1733	1	6-26:2040
1734	1	14-34:1939
1735	1	17-38:1933
1736	2	10-25:2484 95-110:1787
1737	2	5-21:1862 109-125:2714
1738	1	1-21:1881
1739	1	393-414:2402
1740	1	3-21:2189
1741	3	15-33:3158 109-123:2786 364-380:2232
1742	1	564-580:1725
1743	2	40-60:2740 169-189:3006
1744	1	214-228:2582
1745	1	3-19:2099
1746	2	4-22:2753 247-263:2558
1747	1	43-60:2160
1748	1	195-217:2199
1749	4	304-319:1707 427-447:1983 547-562:2389 657-683:2065
1750	5	2-17:2136 753-768:2335 810-831:1715 906-928:1710 1009-1031:2357
1751	1	337-352:2864
1752	1	1-24:3347
1753	1	621-644:2150
1754	1	88-106:2133
1755	4	67-89:2584 130-148:1837 239-261:1922 334-349:2621
1756	1	20-35:1702
1757	1	471-488:1852
1758	2	8-25:2374 247-265:1721
1759	1	10-26:2268
1760	2	37-55:2350 79-102:1867
1761	1	3-16:1751
1762	4	10-29:2708 158-183:2332 413-430:1706 514-532:2104
1763	1	531-558:2381
1764	1	1-15:2095
1765	1	3-20:1765
1766	1	3-27:2376
1767	1	11-39:1868
1768	1	150-169:1896
1769	1	91-112:2655
1770	1	347-362:1792

Table 8  
1134

SEQ ID NO:	Number of Transmembrane Domains	For each Transmembrane Domain, its Transmembrane Domain Position in SEQ ID NO: and its TM Pred Score
1771	1	717-737:3080
1772	1	9-24:1787
1773	2	1-16:1977 85-110:2971
1774	3	492-510:2266 517-532:1752 675-696:1808
1775	3	16-33:2461 94-113:2485 221-247:2232
1776	1	24-39:2149
1777	1	108-123:2281
1778	2	23-42:3401 179-201:3489
1779	1	296-311:1881
1780	1	280-303:2645
1781	1	22-37:2170
1782	1	313-332:1963
1783	7	243-261:2338 288-311:1992 401-429:1705 502-519:1819 556-568:1888 597-614:1780 675-690:2102
1784	1	1433-1457:3108
1785	1	390-403:1855
1786	2	103-122:1737 738-756:2220
1787	2	116-131:2359 173-193:1746
1788	1	128-144:2443
1789	1	73-97:1982
1790	1	103-124:2547
1791	1	53-69:2536
1792	3	84-99:2680 1083-1098:1729 1147-1167:2267
1793	1	59-74:2644
1794	2	24-39:2073 69-92:1922
1795	1	59-78:2490
1796	2	1101-1121:2566 1224-1244:2000
1797	3	54-73:1789 113-131:1860 182-195:2408
1798	2	76-92:1943 135-155:3247
1799	1	141-159:3106
1800	1	248-266:1883
1801	2	413-430:1851 496-521:2037
1802	1	782-811:1739
1803	3	85-108:1832 117-138:1746 166-187:2659
1804	1	183-200:1880
1805	1	549-568:1822
1806	2	121-142:2091 191-210:2580
1807	3	52-67:2292 112-127:1756 215-232:2516
1808	1	99-115:1824
1809	1	24-41:1826
1810	3	710-738:2388 769-792:1779 797-815:1896
1811	1	71-93:2069
1812	1	97-117:1891
1813	3	12-27:2868 38-54:1801 59-89:3042
1814	1	404-422:2347
1815	1	1-19:2068
1816	2	62-84:2012 118-134:1970
1817	1	68-83:2632
1818	2	3-22:2491 47-61:1754
1819	1	1187-1209:3678
1820	6	4-29:1727 87-110:3204 95-129:1851 165-190:2822 214-230:1867 245-261:1911
1821	1	324-342:2300
1822	1	4-24:1911

Table 9  
1135

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No., SEQ ID NO.) *
1	912	1823	2479	784_2167
2	913	1824	2480	790_12732
3	914	1825	2481	787_7952
4	915	1826	2482	787_7952
5	916			
6	917	1827	2483	787_5660
7	918	1828	2484	790_6350
8	919	1829	2485	790_18138
9	920	1830	2486	784_2658
10	921	1831	2487	788_11684
11	922			
12	923	1832	2488	789_1765
13	924	1833	2489	784_6512
14	925	1834	2490	789_4745
15	926	1835	2491	790_14655
16	927	1836	2492	791_2272
17	928	1837	2493	791_1348
18	929	1838	2494	790_23
19	930	1839	2495	784_5039
20	931	1840	2496	787_8518
21	932	1841	2497	784_5125
22	933			
23	934			
24	935	1842	2498	790_26893
25	936	1843	2499	785_3384
26	937	1844	2500	791_2238
27	938	1845	2501	784_7881
28	939			
29	940			
30	941	1846	2502	785_197
31	942	1847	2503	785_836
32	943	1848	2504	784_8990
33	944	1849	2505	792_6041
34	945	1850	2506	792_873
35	946			
36	947	1851	2507	784_4339
37	948			
38	949	1852	2508	792_5188
39	950			
40	951	1853	2509	784_6463
41	952	1854	2510	784_2235
42	953	1855	2511	784_2849
43	954	1856	2512	785_206
44	955	1857	2513	787_6368
45	956	1858	2514	788_5838
46	957	1859	2515	784_7012
47	958			
48	959	1860	2516	785_835
49	960	1861	2517	784_7305
50	961			
51	962	1862	2518	784_8815

Table 9  
1136

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
52	963	1863	2519	784_7195
53	964	1864	2520	787_6723
54	965			
55	966	1865	2521	787_5698
56	967	1866	2522	790_12578
57	968	1867	2523	784_6485
58	969	1868	2524	784_1896
59	970	1869	2525	792_1194
60	971	1870	2526	784_9157
61	972			
62	973	1871	2527	784_1408
63	974	1872	2528	784_2508
64	975	1873	2529	784_1049
65	976			
66	977			
67	978			
68	979			
69	980			
70	981	1874	2530	787_6121
71	982	1875	2531	785_2903
72	983	1876	2532	787_10197
73	984			
74	985	1877	2533	787_10224
75	986	1878	2534	789_5808
76	987			
77	988	1879	2535	784_6287
78	989	1880	2536	785_582
79	990			
80	991			
81	992	1881	2537	790_12952
82	993			
83	994	1882	2538	790_13887
84	995	1883	2539	784_7925
85	996			
86	997			
87	998	1884	2540	790_11010
88	999	1885	2541	790_17371
89	1000			
90	1001			
91	1002	1886	2542	784_1560
92	1003	1887	2543	784_1260
93	1004	1888	2544	790_21336
94	1005	1889	2545	784_7912
95	1006			
96	1007	1890	2546	784_8528
97	1008	1891	2547	784_8105
98	1009	1892	2548	787_5314
99	1010	1893	2549	784_7469
100	1011	1894	2550	790_11802
101	1012	1895	2551	787_5599
102	1013	1896	2552	784_5534



Table 9  
1137

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
103	1014			
104	1015	1897	2553	791_3043
105	1016	1898	2554	784_8916
106	1017	1899	2555	784_1271
107	1018			
108	1019			
109	1020	1900	2556	784_6527
110	1021			
111	1022	1901	2557	790_14183
112	1023	1902	2558	791_1570
113	1024	1903	2559	787_2154
114	1025	1904	2560	787_2849
115	1026			
116	1027			
117	1028			
118	1029	1905	2561	790_23824
119	1030	1906	2562	787_10284
120	1031	1907	2563	790_13062
121	1032			
122	1033			
123	1034			
124	1035	1908	2564	787_3576
125	1036			
126	1037			
127	1038	1909	2565	789_2490
128	1039	1910	2566	784_2340
129	1040	1911	2567	784_3449
130	1041	1912	2568	787_9834
131	1042	1913	2569	787_7563
132	1043			
133	1044	1914	2570	784_969
134	1045	1915	2571	784_9071
135	1046	1916	2572	785_1259
136	1047	1917	2573	790_19506
137	1048	1918	2574	784_3978
138	1049	1919	2575	784_3848
139	1050	1920	2576	789_6174
140	1051	1921	2577	785_1465
141	1052	1922	2578	787_7763
142	1053	1923	2579	790_25889
143	1054			
144	1055	1924	2580	784_4674
145	1056	1925	2581	787_8960
146	1057	1926	2582	787_2258
147	1058	1927	2583	787_2290
148	1059	1928	2584	784_8266
149	1060	1929	2585	784_1397
150	1061	1930	2586	784_8164
151	1062	1931	2587	785_1003
152	1063			
153	1064			

Table 9  
1138

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
154	1065	1932	2588	784_2498
155	1066	1933	2589	787_5189
156	1067			
157	1068	1934	2590	787_10359
158	1069			
159	1070			
160	1071	1935	2591	790_17261
161	1072	1936	2592	784_9629
162	1073	1937	2593	784_3405
163	1074	1938	2594	784_3405
164	1075	1939	2595	790_4043
165	1076	1940	2596	787_3900
166	1077	1941	2597	787_9753
167	1078	1942	2598	787_4766
168	1079	1943	2599	789_3521
169	1080	1944	2600	784_8097
170	1081			
171	1082	1945	2601	784_735
172	1083	1946	2602	784_4418
173	1084	1947	2603	784_4418
174	1085			
175	1086			
176	1087	1948	2604	784_1006
177	1088	1949	2605	784_1528
178	1089			
179	1090			
180	1091	1950	2606	791_2053
181	1092			
182	1093	1951	2607	788_6860
183	1094	1952	2608	785_1010
184	1095	1953	2609	784_6854
185	1096	1954	2610	784_8679
186	1097	1955	2611	790_19249
187	1098	1956	2612	784_5566
188	1099	1957	2613	790_3027
189	1100	1958	2614	790_3027
190	1101	1959	2615	790_3027
191	1102			
192	1103	1960	2616	791_2652
193	1104	1961	2617	784_7116
194	1105	1962	2618	789_1658
195	1106	1963	2619	787_7673
196	1107	1964	2620	790_3240
197	1108			
198	1109	1965	2621	784_6361
199	1110	1966	2622	784_297
200	1111	1967	2623	790_13930
201	1112			
202	1113	1968	2624	784_8933
203	1114			
204	1115			

Table 9  
1139

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No., SEQ ID NO.) *
205	1116			
206	1117			
207	1118	1969	2625	784_7603
208	1119	1970	2626	787_5453
209	1120	1971	2627	789_6377
210	1121	1972	2628	784_7446
211	1122	1973	2629	784_5422
212	1123	1974	2630	784_2530
213	1124	1975	2631	787_7257
214	1125	1976	2632	784_5473
215	1126	1977	2633	784_1793
216	1127	1978	2634	784_10027
217	1128			
218	1129	1979	2635	787_2898
219	1130	1980	2636	788_9208
220	1131	1981	2637	787_2305
221	1132	1982	2638	784_8556
222	1133			
223	1134	1983	2639	787_5766
224	1135	1984	2640	790_4531
225	1136	1985	2641	784_6708
226	1137	1986	2642	790_19316
227	1138	1987	2643	784_1784
228	1139	1988	2644	784_1698
229	1140	1989	2645	787_1368
230	1141	1990	2646	789_6192
231	1142	1991	2647	784_4498
232	1143			
233	1144	1992	2648	789_6042
234	1145			
235	1146	1993	2649	790_4461
236	1147			
237	1148	1994	2650	784_2675
238	1149	1995	2651	789_4591
239	1150	1996	2652	789_4591
240	1151	1997	2653	790_13145
241	1152			
242	1153	1998	2654	791_4932
243	1154			
244	1155	1999	2655	784_10141
245	1156	2000	2656	784_10225
246	1157	2001	2657	784_7722
247	1158			
248	1159			
249	1160	2002	2658	787_3228
250	1161	2003	2659	785_3642
251	1162	2004	2660	784_6469
252	1163	2005	2661	787_6106
253	1164	2006	2662	791_2793
254	1165			
255	1166	2007	2663	784_3318

Table 9  
1140

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
256	1167			
257	1168	2008	2664	789_6270
258	1169	2009	2665	788_13058
259	1170	2010	2666	787_5541
260	1171			
261	1172	2011	2667	784_7891
262	1173	2012	2668	784_4405
263	1174	2013	2669	784_7585
264	1175	2014	2670	784_7585
265	1176	2015	2671	785_3390
266	1177	2016	2672	789_6254
267	1178	2017	2673	785_3073
268	1179	2018	2674	790_16443
269	1180			
270	1181	2019	2675	790_21097
271	1182	2020	2676	787_8948
272	1183	2021	2677	784_8624
273	1184			
274	1185	2022	2678	785_2309
275	1186			
276	1187	2023	2679	784_1834
277	1188	2024	2680	787_8921
278	1189	2025	2681	784_7125
279	1190	2026	2682	785_2903
280	1191	2027	2683	785_2903
281	1192			
282	1193			
283	1194	2028	2684	791_2389
284	1195			
285	1196	2029	2685	784_7647
286	1197	2030	2686	784_3528
287	1198	2031	2687	784_5750
288	1199			
289	1200	2032	2688	789_3496
290	1201			
291	1202			
292	1203	2033	2689	790_16699
293	1204	2034	2690	785_3735
294	1205	2035	2691	788_12224
295	1206	2036	2692	787_6084
296	1207			
297	1208	2037	2693	784_330
298	1209	2038	2694	791_2972
299	1210	2039	2695	790_24002
300	1211			
301	1212	2040	2696	784_528
302	1213	2041	2697	787_1611
303	1214	2042	2698	785_1187
304	1215	2043	2699	784_8087
305	1216	2044	2700	789_4427
306	1217	2045	2701	789_4427

Table 9  
1141

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
307	1218			
308	1219	2046	2702	785_2691
309	1220	2047	2703	785_1940
310	1221	2048	2704	784_5176
311	1222	2049	2705	784_1451
312	1223	2050	2706	789_5236
313	1224	2051	2707	787_5175
314	1225	2052	2708	784_8301
315	1226	2053	2709	785_3698
316	1227	2054	2710	785_2838
317	1228			
318	1229			
319	1230	2055	2711	784_10105
320	1231	2056	2712	784_4974
321	1232	2057	2713	785_3025
322	1233	2058	2714	785_3028
323	1234	2059	2715	785_3031
324	1235	2060	2716	785_2106
325	1236	2061	2717	785_3032
326	1237	2062	2718	792_7466
327	1238	2063	2719	785_2107
328	1239	2064	2720	785_3035
329	1240	2065	2721	785_3037
330	1241	2066	2722	785_3038
331	1242	2067	2723	785_3038
332	1243	2068	2724	785_2108
333	1244	2069	2725	785_3043
334	1245	2070	2726	785_3045
335	1246	2071	2727	785_2114
336	1247	2072	2728	785_483
337	1248	2073	2729	785_609
338	1249	2074	2730	785_1502
339	1250	2075	2731	785_2157
340	1251	2076	2732	785_3106
341	1252	2077	2733	785_3508
342	1253			
343	1254	2078	2734	785_3114
344	1255	2079	2735	785_3115
345	1256	2080	2736	785_2161
346	1257			
347	1258	2081	2737	785_3123
348	1259	2082	2738	785_854
349	1260	2083	2739	785_1504
350	1261	2084	2740	790_3585
351	1262			
352	1263	2085	2741	784_1062
353	1264	2086	2742	784_9616
354	1265			
355	1266	2087	2743	788_604
356	1267	2088	2744	784_9474
357	1268	2089	2745	784_9474

Table 9  
1142

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
358	1269			
359	1270			
360	1271	2090	2746	788_6828
361	1272	2091	2747	788_6828
362	1273	2092	2748	784_8575
363	1274	2093	2749	790_21803
364	1275	2094	2750	784_5329
365	1276	2095	2751	784_10122
366	1277	2096	2752	784_3644
367	1278			
368	1279	2097	2753	787_2863
369	1280	2098	2754	784_854
370	1281			
371	1282			
372	1283	2099	2755	787_3968
373	1284	2100	2756	785_2864
374	1285	2101	2757	785_2864
375	1286			
376	1287	2102	2758	784_4739
377	1288			
378	1289	2103	2759	791_771
379	1290	2104	2760	792_4497
380	1291	2105	2761	784_4833
381	1292	2106	2762	785_3195
382	1293	2107	2763	784_6243
383	1294	2108	2764	784_1134
384	1295	2109	2765	784_10176
385	1296	2110	2766	784_3781
386	1297	2111	2767	785_2441
387	1298	2112	2768	787_4817
388	1299	2113	2769	784_3255
389	1300			
390	1301	2114	2770	784_9436
391	1302	2115	2771	784_5232
392	1303	2116	2772	784_4604
393	1304			
394	1305	2117	2773	784_7986
395	1306	2118	2774	787_10365
396	1307	2119	2775	784_4897
397	1308			
398	1309	2120	2776	787_5541
399	1310	2121	2777	787_7655
400	1311	2122	2778	784_1002
401	1312	2123	2779	787_8365
402	1313			
403	1314	2124	2780	784_9956
404	1315	2125	2781	787_7927
405	1316	2126	2782	787_10138
406	1317	2127	2783	787_5878
407	1318	2128	2784	784_6962
408	1319	2129	2785	784_1341

Table 9  
1143

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
409	1320	2130	2786	787_894
410	1321			
411	1322	2131	2787	784_5866
412	1323	2132	2788	784_5734
413	1324			
414	1325			
415	1326	2133	2789	788_11985
416	1327			
417	1328	2134	2790	791_1605
418	1329			
419	1330	2135	2791	784_3232
420	1331	2136	2792	784_1670
421	1332			
422	1333	2137	2793	784_2377
423	1334	2138	2794	788_2631
424	1335	2139	2795	784_8666
425	1336	2140	2796	785_1109
426	1337	2141	2797	785_1109
427	1338	2142	2798	784_6176
428	1339	2143	2799	785_1465
429	1340	2144	2800	785_1465
430	1341	2145	2801	784_5369
431	1342	2146	2802	787_8763
432	1343	2147	2803	784_5253
433	1344	2148	2804	789_2568
434	1345	2149	2805	790_3055
435	1346	2150	2806	790_10356
436	1347	2151	2807	785_2811
437	1348	2152	2808	784_2496
438	1349	2153	2809	784_2530
439	1350	2154	2810	784_2168
440	1351	2155	2811	784_7505
441	1352	2156	2812	784_6670
442	1353	2157	2813	787_2168
443	1354	2158	2814	784_3022
444	1355	2159	2815	790_1217
445	1356	2160	2816	784_5473
446	1357	2161	2817	784_7905
447	1358			
448	1359	2162	2818	784_9436
449	1360	2163	2819	787_2898
450	1361			
451	1362			
452	1363	2164	2820	784_4625
453	1364	2165	2821	785_2622
454	1365			
455	1366	2166	2822	787_2386
456	1367	2167	2823	788_9391
457	1368	2168	2824	784_10104
458	1369	2169	2825	784_10104
459	1370	2170	2826	787_10366

Table 9  
1144

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
460	1371	2171	2827	784_6453
461	1372	2172	2828	785_1337
462	1373	2173	2829	787_5778
463	1374	2174	2830	784_5025
464	1375			
465	1376	2175	2831	785_739
466	1377	2176	2832	787_1269
467	1378			
468	1379			
469	1380			
470	1381	2177	2833	785_326
471	1382	2178	2834	784_8341
472	1383	2179	2835	787_757
473	1384			
474	1385			
475	1386	2180	2836	787_8454
476	1387	2181	2837	790_8424
477	1388			
478	1389	2182	2838	787_9580
479	1390	2183	2839	784_6906
480	1391	2184	2840	787_1411
481	1392	2185	2841	790_10999
482	1393			
483	1394			
484	1395			
485	1396			
486	1397	2186	2842	787_2916
487	1398	2187	2843	787_5986
488	1399	2188	2844	785_2979
489	1400	2189	2845	785_2105
490	1401			
491	1402			
492	1403	2190	2846	785_3046
493	1404	2191	2847	788_9546
494	1405	2192	2848	785_2158
495	1406			
496	1407	2193	2849	784_7549
497	1408	2194	2850	784_1066
498	1409	2195	2851	784_2673
499	1410	2196	2852	785_563
500	1411			
501	1412	2197	2853	789_630
502	1413	2198	2854	785_259
503	1414	2199	2855	784_5637
504	1415	2200	2856	790_7396
505	1416			
506	1417	2201	2857	790_11548
507	1418			
508	1419	2202	2858	784_6740
509	1420			
510	1421	2203	2859	785_1007



Table 9  
1145

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
511	1422			
512	1423			
513	1424	2204	2860	784_3641
514	1425			
515	1426	2205	2861	789_6066
516	1427	2206	2862	788_13582
517	1428	2207	2863	784_9762
518	1429	2208	2864	787_2052
519	1430	2209	2865	790_3055
520	1431			
521	1432	2210	2866	790_23800
522	1433	2211	2867	790_23800
523	1434			
524	1435	2212	2868	787_2548
525	1436			
526	1437	2213	2869	784_2565
527	1438	2214	2870	784_544
528	1439	2215	2871	784_544
529	1440			
530	1441	2216	2872	789_1844
531	1442	2217	2873	787_2733
532	1443	2218	2874	785_493
533	1444	2219	2875	790_18203
534	1445			
535	1446	2220	2876	784_3336
536	1447	2221	2877	790_777
537	1448	2222	2878	787_4220
538	1449	2223	2879	790_26438
539	1450	2224	2880	790_13586
540	1451	2225	2881	784_270
541	1452			
542	1453	2226	2882	790_14176
543	1454	2227	2883	784_3543
544	1455	2228	2884	784_3575
545	1456	2229	2885	787_5638
546	1457	2230	2886	784_3826
547	1458	2231	2887	787_6059
548	1459	2232	2888	787_2585
549	1460			
550	1461			
551	1462	2233	2889	787_2965
552	1463			
553	1464	2234	2890	790_13050
554	1465			
555	1466	2235	2891	784_3374
556	1467	2236	2892	787_8240
557	1468	2237	2893	784_4398
558	1469			
559	1470	2238	2894	788_10891
560	1471	2239	2895	787_9805
561	1472			

Table 9  
1146

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
562	1473			
563	1474	2240	2896	787_7094
564	1475	2241	2897	787_5209
565	1476	2242	2898	784_582
566	1477	2243	2899	790_10111
567	1478	2244	2900	790_13563
568	1479			
569	1480	2245	2901	787_8234
570	1481			
571	1482	2246	2902	790_16394
572	1483			
573	1484	2247	2903	790_18452
574	1485	2248	2904	790_211
575	1486	2249	2905	790_14932
576	1487			
577	1488	2250	2906	787_7498
578	1489	2251	2907	787_610
579	1490	2252	2908	790_19096
580	1491	2253	2909	784_7363
581	1492	2254	2910	792_8355
582	1493	2255	2911	787_1179
583	1494	2256	2912	790_6822
584	1495	2257	2913	784_8670
585	1496	2258	2914	784_2469
586	1497			
587	1498	2259	2915	790_15720
588	1499	2260	2916	790_11300
589	1500	2261	2917	784_190
590	1501			
591	1502	2262	2918	789_6245
592	1503	2263	2919	787_7385
593	1504	2264	2920	790_281
594	1505	2265	2921	784_6192
595	1506	2266	2922	790_770
596	1507	2267	2923	784_3364
597	1508	2268	2924	784_5424
598	1509			
599	1510	2269	2925	790_10118
600	1511			
601	1512			
602	1513	2270	2926	785_2555
603	1514			
604	1515			
605	1516	2271	2927	790_20185
606	1517	2272	2928	784_333
607	1518	2273	2929	792_3164
608	1519			
609	1520			
610	1521	2274	2930	784_3455
611	1522	2275	2931	784_506
612	1523	2276	2932	788_7179

Table 9  
1147

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
613	1524			
614	1525			
615	1526	2277	2933	790_13445
616	1527			
617	1528	2278	2934	784_6181
618	1529	2279	2935	784_4453
619	1530			
620	1531	2280	2936	784_8140
621	1532			
622	1533	2281	2937	784_3441
623	1534	2282	2938	787_6194
624	1535	2283	2939	784_1407
625	1536			
626	1537	2284	2940	790_8886
627	1538			
628	1539	2285	2941	785_2555
629	1540	2286	2942	784_688
630	1541	2287	2943	784_8302
631	1542	2288	2944	787_2581
632	1543	2289	2945	790_19616
633	1544	2290	2946	784_4663
634	1545	2291	2947	790_9138
635	1546			
636	1547	2292	2948	784_4743
637	1548	2293	2949	790_12541
638	1549			
639	1550			
640	1551	2294	2950	784_4128
641	1552			
642	1553	2295	2951	790_19867
643	1554	2296	2952	784_7061
644	1555	2297	2953	785_1399
645	1556	2298	2954	784_5565
646	1557	2299	2955	790_10780
647	1558			
648	1559			
649	1560	2300	2956	784_2148
650	1561	2301	2957	790_24174
651	1562	2302	2958	784_4307
652	1563	2303	2959	790_20063
653	1564			
654	1565	2304	2960	784_7654
655	1566	2305	2961	784_7454
656	1567	2306	2962	790_8324
657	1568	2307	2963	784_5860
658	1569			
659	1570	2308	2964	788_11718
660	1571	2309	2965	784_2105
661	1572			
662	1573	2310	2966	785_1924
663	1574	2311	2967	784_8098

Table 9  
1148

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
664	1575	2312	2968	784_2320
665	1576	2313	2969	784_7695
666	1577	2314	2970	784_6847
667	1578			
668	1579			
669	1580	2315	2971	784_8098
670	1581	2316	2972	784_4781
671	1582			
672	1583	2317	2973	792_8335
673	1584			
674	1585	2318	2974	789_2488
675	1586	2319	2975	785_2433
676	1587			
677	1588	2320	2976	787_3115
678	1589			
679	1590	2321	2977	787_5209
680	1591	2322	2978	787_4223
681	1592	2323	2979	784_5228
682	1593	2324	2980	787_5094
683	1594	2325	2981	790_10116
684	1595			
685	1596	2326	2982	785_1351
686	1597			
687	1598			
688	1599	2327	2983	784_4288
689	1600	2328	2984	784_4816
690	1601	2329	2985	791_3535
691	1602	2330	2986	784_4322
692	1603	2331	2987	785_2272
693	1604	2332	2988	790_12959
694	1605	2333	2989	784_4642
695	1606	2334	2990	784_3534
696	1607			
697	1608			
698	1609			
699	1610	2335	2991	784_2904
700	1611	2336	2992	784_8547
701	1612			
702	1613			
703	1614	2337	2993	785_2907
704	1615			
705	1616	2338	2994	789_3856
706	1617	2339	2995	785_2043
707	1618	2340	2996	784_4327
708	1619	2341	2997	787_10225
709	1620			
710	1621	2342	2998	790_24136
711	1622	2343	2999	785_172
712	1623	2344	3000	790_24958
713	1624	2345	3001	790_17848
714	1625	2346	3002	790_10920

Table 9  
1149

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
715	1626	2347	3003	790_24477
716	1627	2348	3004	790_21024
717	1628	2349	3005	790_3053
718	1629			
719	1630	2350	3006	790_13392
720	1631	2351	3007	790_28822
721	1632	2352	3008	790_17805
722	1633	2353	3009	784_1991
723	1634	2354	3010	790_12778
724	1635	2355	3011	784_5899
725	1636			
726	1637	2356	3012	784_4851
727	1638	2357	3013	784_6070
728	1639	2358	3014	784_9615
729	1640	2359	3015	790_23719
730	1641			
731	1642	2360	3016	784_7918
732	1643	2361	3017	790_16176
733	1644			
734	1645	2362	3018	784_4042
735	1646			
736	1647	2363	3019	790_4484
737	1648			
738	1649	2364	3020	790_16690
739	1650	2365	3021	784_7703
740	1651	2366	3022	785_2626
741	1652			
742	1653	2367	3023	784_1423
743	1654			
744	1655			
745	1656	2368	3024	784_8178
746	1657	2369	3025	784_6726
747	1658	2370	3026	792_7162
748	1659	2371	3027	784_2948
749	1660			
750	1661	2372	3028	784_2988
751	1662			
752	1663	2373	3029	784_10047
753	1664			
754	1665	2374	3030	787_3415
755	1666			
756	1667	2375	3031	787_10007
757	1668			
758	1669			
759	1670	2376	3032	790_12779
760	1671			
761	1672	2377	3033	785_3699
762	1673	2378	3034	784_4026
763	1674	2379	3035	790_14323
764	1675			
765	1676			

Table 9  
1150

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
766	1677	2380	3036	792_2616
767	1678	2381	3037	787_6570
768	1679	2382	3038	784_2590
769	1680	2383	3039	784_5216
770	1681			
771	1682	2384	3040	790_12514
772	1683			
773	1684			
774	1685	2385	3041	785_3366
775	1686	2386	3042	790_12446
776	1687			
777	1688			
778	1689			
779	1690			
780	1691	2387	3043	784_6708
781	1692	2388	3044	784_852
782	1693			
783	1694			
784	1695			
785	1696	2389	3045	784_9460
786	1697	2390	3046	787_7895
787	1698	2391	3047	784_6226
788	1699	2392	3048	789_5263
789	1700	2393	3049	784_4933
790	1701			
791	1702			
792	1703			
793	1704	2394	3050	784_7176
794	1705			
795	1706	2395	3051	784_3074
796	1707			
797	1708	2396	3052	785_1642
798	1709			
799	1710	2397	3053	790_12671
800	1711			
801	1712			
802	1713	2398	3054	784_1127
803	1714			
804	1715			
805	1716	2399	3055	790_14423
806	1717	2400	3056	790_28166
807	1718			
808	1719			
809	1720	2401	3057	784_5946
810	1721	2402	3058	784_4850
811	1722	2403	3059	784_744
812	1723	2404	3060	790_13151
813	1724			
814	1725	2405	3061	787_7172
815	1726	2406	3062	790_6384
816	1727			

Table 9  
1151

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
817	1728			
818	1729	2407	3063	785_3645
819	1730	2408	3064	784_9630
820	1731			
821	1732	2409	3065	784_10027
822	1733	2410	3066	784_10083
823	1734	2411	3067	787_2031
824	1735	2412	3068	784_4592
825	1736			
826	1737	2413	3069	790_3308
827	1738			
828	1739	2414	3070	784_4360
829	1740	2415	3071	785_2472
830	1741	2416	3072	784_2060
831	1742	2417	3073	784_3549
832	1743			
833	1744	2418	3074	784_1872
834	1745	2419	3075	784_7925
835	1746	2420	3076	784_5564
836	1747			
837	1748			
838	1749	2421	3077	785_1472
839	1750	2422	3078	784_2232
840	1751	2423	3079	790_24132
841	1752	2424	3080	790_6689
842	1753	2425	3081	787_3699
843	1754	2426	3082	785_2955
844	1755	2427	3083	784_4324
845	1756	2428	3084	784_9382
846	1757	2429	3085	787_7087
847	1758			
848	1759	2430	3086	784_4912
849	1760			
850	1761	2431	3087	790_17236
851	1762	2432	3088	790_1914
852	1763	2433	3089	784_466
853	1764	2434	3090	790_11164
854	1765			
855	1766	2435	3091	787_8100
856	1767	2436	3092	784_9062
857	1768			
858	1769	2437	3093	784_4141
859	1770			
860	1771	2438	3094	784_1365
861	1772	2439	3095	784_6642
862	1773			
863	1774	2440	3096	789_3442
864	1775	2441	3097	784_10198
865	1776	2442	3098	790_11264
866	1777			
867	1778	2443	3099	788_13640

Table 9  
1152

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Identification of Priority Application that contig nucleotide sequence was filed (Attorney Docket No. SEQ ID NO.) *
868	1779			
869	1780	2444	3100	785_238
870	1781	2445	3101	787_10243
871	1782	2446	3102	784_8555
872	1783	2447	3103	789_5245
873	1784	2448	3104	784_4968
874	1785	2449	3105	789_5631
875	1786	2450	3106	784_6373
876	1787	2451	3107	784_7992
877	1788	2452	3108	787_6002
878	1789	2453	3109	784_1503
879	1790	2454	3110	790_10504
880	1791	2455	3111	790_18438
881	1792	2456	3112	790_13790
882	1793	2457	3113	790_21890
883	1794	2458	3114	790_17754
884	1795			
885	1796	2459	3115	785_1016
886	1797	2460	3116	785_14
887	1798	2461	3117	790_18472
888	1799	2462	3118	784_8233
889	1800	2463	3119	784_177
890	1801	2464	3120	784_3522
891	1802	2465	3121	784_10067
892	1803	2466	3122	784_7868
893	1804			
894	1805	2467	3123	787_7085
895	1806	2468	3124	784_3739
896	1807	2469	3125	784_4168
897	1808	2470	3126	784_6642
898	1809			
899	1810	2471	3127	789_6290
900	1811	2472	3128	791_2423
901	1812			
902	1813			
903	1814			
904	1815			
905	1816	2473	3129	790_10106
906	1817	2474	3130	787_3075
907	1818	2475	3131	790_17762
908	1819	2476	3132	784_482
909	1820	2477	3133	790_21059
910	1821	2478	3134	788_13847
911	1822			

\*784\_XXX = SEQ ID NO: XXX of Attorney Docket No. 784, US Serial No. 09/488,725 filed 01/21/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.



## Table 9

1153

785\_XXX = SEQ ID NO: XXX of Attorney Docket No. 785, US Serial No. 09/491,404 filed 01/25/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.

787\_XXX = SEQ ID NO: XXX of Attorney Docket No. 787, US Serial No. 09/496,914 filed 02/03/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.

788\_XXX = SEQ ID NO: XXX of Attorney Docket No. 788, US Serial No. 09/515,126 filed 02/28/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.

789\_XXX = SEQ ID NO: XXX of Attorney Docket No. 789, US Serial No. 09/519,705 filed 03/07/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.

790\_XXX = SEQ ID NO: XXX of Attorney Docket No. 790, US Serial No. 09/540,217 filed 03/31/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.

791\_XXX = SEQ ID NO: XXX of Attorney Docket No. 791, US Serial No. 09/552,929 filed 04/18/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.

792\_XXX = SEQ ID NO: XXX of Attorney Docket No. 792, US Serial No. 09/577,408 filed 05/18/2000, the entire disclosure of which, including sequence listing, is incorporated herein by reference.

Table 10  
1154

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/339,739
1	912	1
2	913	2
3	914	3
4	915	4
5	916	5
6	917	6
7	918	7
8	919	8
9	920	9
10	921	10
11	922	11
12	923	12
13	924	13
14	925	14
15	926	15
16	927	16
17	928	17
18	929	18
19	930	19
20	931	20
21	932	21
22	933	22
23	934	23
24	935	24
25	936	25
26	937	26
27	938	27
28	939	28
29	940	29
30	941	30
31	942	31
32	943	32
33	944	33
34	945	34
35	946	35
36	947	36
37	948	37
38	949	38
39	950	39
40	951	40
41	952	41
42	953	42
43	954	43
44	955	44
45	956	45
46	957	46
47	958	47
48	959	48
49	960	49
50	961	50
51	962	51
52	963	52

Table 10  
1155

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/339,739
53	964	53
54	965	54
55	966	55
56	967	56
57	968	57
58	969	58
59	970	59
60	971	60
61	972	61
62	973	62
63	974	63
64	975	64
65	976	65
66	977	66
67	978	67
68	979	68
69	980	69
70	981	70
71	982	71
72	983	72
73	984	73
74	985	74
75	986	75
76	987	76
77	988	77
78	989	78
79	990	79
80	991	80
81	992	81
82	993	82
83	994	83
84	995	84
85	996	85
86	997	86
87	998	87
88	999	88
89	1000	89
90	1001	90
91	1002	91
92	1003	92
93	1004	93
94	1005	94
95	1006	95
96	1007	96
97	1008	97
98	1009	98
99	1010	99
100	1011	100
101	1012	101
102	1013	102
103	1014	103
104	1015	104

Table 10  
1156

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/339,739
105	1016	105
106	1017	106
107	1018	107
108	1019	108
109	1020	109
110	1021	110
111	1022	111
112	1023	112
113	1024	113
114	1025	114
115	1026	115
116	1027	116
117	1028	117
118	1029	118
119	1030	119
120	1031	120
121	1032	121
122	1033	122
123	1034	123
124	1035	124
125	1036	125

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/339,453
126	1037	1
127	1038	2
128	1039	3
129	1040	4
130	1041	5
131	1042	7
132	1043	8
133	1044	9
134	1045	10
135	1046	11
136	1047	12
137	1048	13
138	1049	14
139	1050	15
140	1051	16
141	1052	17
142	1053	18
143	1054	19
144	1055	20
145	1056	21
146	1057	22
147	1058	23
148	1059	24
149	1060	25
150	1061	26
151	1062	27

Table 10  
1157

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/339,453
152	1063	28
153	1064	29
154	1065	30
155	1066	31
156	1067	32
157	1068	33
158	1069	34
159	1070	35
160	1071	36
161	1072	37
162	1073	38
163	1074	39
164	1075	40
165	1076	41
166	1077	42
167	1078	43
168	1079	44
169	1080	45
170	1081	46
171	1082	47
172	1083	48
173	1084	49
174	1085	50
175	1086	51
176	1087	52
177	1088	53
178	1089	54
179	1090	55
180	1091	56
181	1092	57
182	1093	58
183	1094	59
184	1095	60
185	1096	61
186	1097	62
187	1098	63
188	1099	64
189	1100	65
190	1101	66
191	1102	67
192	1103	68
193	1104	69
194	1105	70
195	1106	71
196	1107	72
197	1108	73
198	1109	74
199	1110	75
200	1111	76
201	1112	77
202	1113	78
203	1114	79

Table 10  
1158

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/339,453
204	1115	80
205	1116	81
206	1117	82
207	1118	83
208	1119	84
209	1120	85
210	1121	86
211	1122	87
212	1123	88
213	1124	89
214	1125	90
215	1126	91
216	1127	92
217	1128	93
218	1129	94
219	1130	95
220	1131	96
221	1132	97
222	1133	98
223	1134	99
224	1135	100
225	1136	101
226	1137	102
227	1138	103
228	1139	104
229	1140	105
230	1141	106
231	1142	107
232	1143	108
233	1144	109
234	1145	110
235	1146	111
236	1147	112
237	1148	113
238	1149	114
239	1150	115
240	1151	116
241	1152	117
242	1153	118
243	1154	119
244	1155	120
245	1156	121
246	1157	122
247	1158	123
248	1159	124

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,384
249	1160	3
250	1161	4

Table 10  
1159

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,384
251	1162	5
252	1163	6
253	1164	8
254	1165	12
255	1166	14
256	1167	15
257	1168	17
258	1169	18
259	1170	19
260	1171	20
261	1172	21
262	1173	22
263	1174	23
264	1175	24
265	1176	25
266	1177	26
267	1178	27
268	1179	28
269	1180	29
270	1181	30
271	1182	31
272	1183	32
273	1184	33
274	1185	34
275	1186	35
276	1187	36
277	1188	37
278	1189	38
279	1190	39
280	1191	40
281	1192	41
282	1193	42
283	1194	43
284	1195	44
285	1196	45
286	1197	46
287	1198	47
288	1199	48
289	1200	49
290	1201	50
291	1202	51
292	1203	52
293	1204	53
294	1205	54
295	1206	55
296	1207	56
297	1208	57
298	1209	58
299	1210	59
300	1211	60
301	1212	61
302	1213	62

Table 10  
1160

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,384
303	1214	63
304	1215	64
305	1216	65
306	1217	66
307	1218	67
308	1219	68
309	1220	69
310	1221	70
311	1222	71
312	1223	72
313	1224	73
314	1225	74
315	1226	75
316	1227	76
317	1228	77
318	1229	78
319	1230	79
320	1231	80
321	1232	81
322	1233	82
323	1234	83
324	1235	84
325	1236	85
326	1237	86
327	1238	87
328	1239	88
329	1240	89
330	1241	90
331	1242	91
332	1243	92
333	1244	93
334	1245	94
335	1246	95
336	1247	96
337	1248	97
338	1249	98
339	1250	99
340	1251	100
341	1252	101
342	1253	102
343	1254	103
344	1255	104
345	1256	105
346	1257	106
347	1258	107
348	1259	108
349	1260	109
350	1261	110
351	1262	111
352	1263	112
353	1264	113
354	1265	114



Table 10  
1161

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,384
355	1266	115
356	1267	116
357	1268	117
358	1269	118
359	1270	119
360	1271	120
361	1272	121
362	1273	122
363	1274	123
364	1275	124
365	1276	125
366	1277	126
367	1278	127
368	1279	128
369	1280	129
370	1281	130
371	1282	131
372	1283	132
373	1284	133
374	1285	134
375	1286	135
376	1287	136
377	1288	137
378	1289	138
379	1290	139
380	1291	140
381	1292	141
382	1293	142
383	1294	143
384	1295	144
385	1296	145
386	1297	146
387	1298	147
388	1299	148
389	1300	149
390	1301	150
391	1302	151
392	1303	152
393	1304	153
394	1305	154
395	1306	155
396	1307	156
397	1308	157
398	1309	158
399	1310	159
400	1311	160
401	1312	161
402	1313	162
403	1314	163
404	1315	164
405	1316	165
406	1317	166

Table 10  
1162

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,384
407	1318	167
408	1319	169
409	1320	170
410	1321	171
411	1322	172
412	1323	173
413	1324	174
414	1325	175
415	1326	176
416	1327	177
417	1328	178
418	1329	179

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,091
419	1330	2
420	1331	4
421	1332	5
422	1333	6
423	1334	8
424	1335	9
425	1336	10
426	1337	11
427	1338	12
428	1339	14
429	1340	15
430	1341	16
431	1342	17
432	1343	18
433	1344	19
434	1345	20
435	1346	21
436	1347	22
437	1348	24
438	1349	25
439	1350	26
440	1351	27
441	1352	28
442	1353	29
443	1354	32
444	1355	33
445	1356	34
446	1357	35
447	1358	36
448	1359	37
449	1360	38
450	1361	39
451	1362	40
452	1363	41
453	1364	42

Table 10  
1163

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,091
454	1365	43
455	1366	44
456	1367	45
457	1368	46
458	1369	47
459	1370	48
460	1371	49
461	1372	50
462	1373	51
463	1374	52
464	1375	53
465	1376	54
466	1377	55
467	1378	56
468	1379	57
469	1380	58
470	1381	59
471	1382	60
472	1383	61
473	1384	62
474	1385	63
475	1386	64
476	1387	65
477	1388	66
478	1389	67
479	1390	68
480	1391	69
481	1392	70
482	1393	71
483	1394	72
484	1395	73
485	1396	74
486	1397	75
487	1398	76
488	1399	77
489	1400	78
490	1401	79
491	1402	80
492	1403	81
493	1404	82
494	1405	83
495	1406	84
496	1407	85
497	1408	86
498	1409	87
499	1410	88
500	1411	89
501	1412	90
502	1413	91
503	1414	92
504	1415	93
505	1416	94

Table 10  
1164

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/365,091
506	1417	95
507	1418	96
508	1419	97
509	1420	98
510	1421	99
511	1422	100
512	1423	101
513	1424	102
514	1425	103
515	1426	104
516	1427	105
517	1428	106
518	1429	107
519	1430	108
520	1431	109
521	1432	110
522	1433	111
523	1434	112
524	1435	113
525	1436	114
526	1437	115
527	1438	116
528	1439	117
529	1440	118
530	1441	119
531	1442	120
532	1443	121
533	1444	122
534	1445	123
535	1446	124
536	1447	125
537	1448	126
538	1449	127
539	1450	128
540	1451	129
541	1452	130
542	1453	131
543	1454	132
544	1455	133
545	1456	135
546	1457	136
547	1458	137
548	1459	138
549	1460	139
550	1461	140
551	1462	141
552	1463	142
553	1464	143
554	1465	144
555	1466	145

Table 10  
1165

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,615
556	1467	2
557	1468	3
558	1469	4
559	1470	5
560	1471	6
561	1472	7
562	1473	8
563	1474	9
564	1475	10
565	1476	11
566	1477	12
567	1478	13
568	1479	14
569	1480	15
570	1481	16
571	1482	17
572	1483	18
573	1484	19
574	1485	20
575	1486	21
576	1487	22
577	1488	23
578	1489	24
579	1490	25
580	1491	26
581	1492	27
582	1493	28
583	1494	29
584	1495	30
585	1496	31
586	1497	32
587	1498	33
588	1499	34
589	1500	35
590	1501	36
591	1502	37
592	1503	38
593	1504	39
594	1505	40
595	1506	41
596	1507	42
597	1508	43
598	1509	44
599	1510	45
600	1511	46
601	1512	47
602	1513	48
603	1514	49
604	1515	50
605	1516	51
606	1517	52
607	1518	53

Table 10  
1166

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,615
608	1519	54
609	1520	55
610	1521	56
611	1522	57
612	1523	58
613	1524	59
614	1525	60
615	1526	61
616	1527	62
617	1528	63
618	1529	64
619	1530	65
620	1531	66
621	1532	67
622	1533	68
623	1534	69
624	1535	70
625	1536	71
626	1537	72
627	1538	73
628	1539	74
629	1540	75
630	1541	76
631	1542	77
632	1543	78
633	1544	79
634	1545	80
635	1546	81
636	1547	82
637	1548	83
638	1549	84
639	1550	85
640	1551	86
641	1552	87
642	1553	88
643	1554	89
644	1555	90
645	1556	91
646	1557	92
647	1558	93
648	1559	94
649	1560	95
650	1561	96
651	1562	97
652	1563	98
653	1564	99
654	1565	100
655	1566	101
656	1567	102
657	1568	103
658	1569	104
659	1570	105

Table 10  
1167

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,615
660	1571	106
661	1572	107
662	1573	108
663	1574	109
664	1575	110
665	1576	111
666	1577	112
667	1578	113
668	1579	114
669	1580	115
670	1581	116
671	1582	117
672	1583	118
673	1584	119
674	1585	120
675	1586	121
676	1587	122
677	1588	123
678	1589	124
679	1590	125
680	1591	126
681	1592	127
682	1593	128
683	1594	129
684	1595	130
685	1596	131
686	1597	132
687	1598	133
688	1599	134
689	1600	135
690	1601	136
691	1602	137
692	1603	138
693	1604	139
694	1605	140
695	1606	141
696	1607	142
697	1608	143
698	1609	144
699	1610	145
700	1611	146
701	1612	147
702	1613	148
703	1614	149
704	1615	150
705	1616	151
706	1617	152
707	1618	153
708	1619	154
709	1620	155
710	1621	157
711	1622	158

Table 10  
1168

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,615
712	1623	159
713	1624	160
714	1625	161
715	1626	162
716	1627	163
717	1628	164
718	1629	165
719	1630	166
720	1631	167
721	1632	168
722	1633	169
723	1634	170
724	1635	171
725	1636	172
726	1637	173
727	1638	174
728	1639	175
729	1640	176
730	1641	177
731	1642	178
732	1643	179
733	1644	180
734	1645	181
735	1646	182
736	1647	183
737	1648	184
738	1649	185
739	1650	186
740	1651	187
741	1652	188
742	1653	189
743	1654	190
744	1655	191
745	1656	192
746	1657	193
747	1658	194
748	1659	195
749	1660	196
750	1661	197
751	1662	198
752	1663	199
753	1664	200
754	1665	201
755	1666	202
756	1667	203
757	1668	204
758	1669	205
759	1670	206
760	1671	207
761	1672	208
762	1673	209
763	1674	210



Table 10  
1169

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,615
764	1675	211
765	1676	212
766	1677	213
767	1678	214
768	1679	215
769	1680	216
770	1681	217
771	1682	218
772	1683	219
773	1684	220
774	1685	221
775	1686	222
776	1687	223
777	1688	224
778	1689	225
779	1690	226
780	1691	227
781	1692	228
782	1693	229
783	1694	230
784	1695	231
785	1696	233
786	1697	234
787	1698	235
788	1699	236
789	1700	237
790	1701	238
791	1702	239
792	1703	240
793	1704	241
794	1705	242
795	1706	243
796	1707	244
797	1708	245
798	1709	246
799	1710	247
800	1711	248
801	1712	249
802	1713	250
803	1714	251
804	1715	252
805	1716	253
806	1717	254
807	1718	255
808	1719	256

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,381
809	1720	1
810	1721	2

Table 10  
1170

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,381
811	1722	3
812	1723	4
813	1724	5
814	1725	6
815	1726	7
816	1727	8
817	1728	9
818	1729	10
819	1730	11
820	1731	12
821	1732	13
822	1733	14
823	1734	15
824	1735	16
825	1736	17
826	1737	18
827	1738	19
828	1739	20
829	1740	21
830	1741	22
831	1742	23
832	1743	24
833	1744	25
834	1745	26
835	1746	27
836	1747	28
837	1748	29
838	1749	30
839	1750	31
840	1751	32
841	1752	33
842	1753	34
843	1754	35
844	1755	36
845	1756	37
846	1757	38
847	1758	39
848	1759	40
849	1760	41
850	1761	42
851	1762	43
852	1763	44
853	1764	45
854	1765	46
855	1766	47
856	1767	48
857	1768	49
858	1769	50
859	1770	51
860	1771	52
861	1772	53
862	1773	54

Table 10  
1171

SEQ ID NO of Full-length Nucleotide Sequence	SEQ ID NO of Full-length Peptide Sequence	SEQ ID NO in Priority Application USSN 60/372,381
863	1774	55
864	1775	56
865	1776	57
866	1777	58
867	1778	59
868	1779	60
869	1780	61
870	1781	62
871	1782	63
872	1783	64
873	1784	65
874	1785	66
875	1786	67
876	1787	68
877	1788	69
878	1789	70
879	1790	71
880	1791	72
881	1792	73
882	1793	74
883	1794	75
884	1795	76
885	1796	77
886	1797	78
887	1798	79
888	1799	80
889	1800	81
890	1801	82
891	1802	83
892	1803	84
893	1804	85
894	1805	86
895	1806	87
896	1807	88
897	1808	89
898	1809	90
899	1810	91
900	1811	92
901	1812	93
902	1813	94
903	1814	95
904	1815	96
905	1816	97
906	1817	98
907	1818	99
908	1819	100
909	1820	101
910	1821	102
911	1822	103

## WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-911.
2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 99% sequence identity with the polynucleotide of claim 1.
4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
6. A vector comprising the polynucleotide of claim 1.
7. An expression vector comprising the polynucleotide of claim 1.
8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
  - (a) a polypeptide encoded by any one of the polynucleotides of claim 1;  
and
  - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-911.

11. A composition comprising the polypeptide of claim 10 and a carrier.
12. An antibody directed against the polypeptide of claim 10.
13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
  - a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
  - b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
  - a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
  - b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
  - c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
  - a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
  - b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and

b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of any of the polynucleotides from SEQ ID NO: 1-911, under conditions sufficient to express the polypeptide in said cell; and

b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of any one of the polypeptides SEQ ID NO: 912-1822.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.

22. A collection of polynucleotides, wherein the collection comprising of at least one of SEQ ID NO: 1-911.

23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.

24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.

25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
26. The collection of claim 22, wherein the collection is provided in a computer-readable format.